

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 522.566 Seconds
(without alignments)
15555.198 Million cell updates/sec

Title: US-10-696-900-1_COPY_1_143
Perfect score: 143
Sequence: 1 ttgccacctccctctctgcg.....aactccatcactaggggtaa 143

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	4718	6	BD242766
2	143	100.0	4718	6	AR562498
3	143	100.0	4718	6	AX753251
4	143	100.0	4718	13	AF063497
5	138.8	97.1	4683	13	AF028704
6	137.2	95.9	4683	6	BD242775
7	137.2	95.9	4683	6	AR562507
8	135.6	94.8	4721	6	CQ972062
9	135.6	94.8	4721	6	CS073592
10	135.6	94.8	4721	6	AX753246
11	135.6	94.8	4721	13	AF513851
12	119	83.2	4718	6	BD242766
13	119	83.2	4718	6	AR562498
14	119	83.2	4718	6	AX753251
15	119	83.2	4718	13	AF063497
16	113.2	79.2	4722	13	AF028705
17	113.2	79.2	4726	6	AX753253
18	113.2	79.2	4726	13	AVU48704

19	112.2	78.5	4722	13	AF028705
20	111.6	78.0	4721	6	CQ972062
21	111.6	78.0	4721	6	CS073592
22	111.6	78.0	4721	6	AX753246
23	111.6	78.0	4721	13	AF513851
24	107.4	75.1	145	6	A46401
25	107.4	75.1	145	13	AA2L1TR1
26	107.4	75.1	145	13	AA2RSPORI
27	107.4	75.1	165	6	AR034135
28	107.4	75.1	165	6	BD218219
29	107.4	75.1	165	6	I16806
30	107.4	75.1	165	6	AR223306
31	107.4	75.1	165	6	AR58389
32	107.4	75.1	165	6	AX106702
33	107.4	75.1	207	6	AX703496
34	107.4	75.1	272	6	BD218218
35	107.4	75.1	272	6	AR223305
36	107.4	75.1	272	6	AR658388
37	107.4	75.1	272	6	AX106701
38	107.4	75.1	2116	13	AA2LEFT
39	107.4	75.1	4675	6	BD094552
40	107.4	75.1	4675	6	AX135805
41	107.4	75.1	4675	6	AX286292
42	107.4	75.1	4675	6	AX753252
43	107.4	75.1	4675	13	AA2CG
44	107.4	75.1	4679	6	AX282480
45	107.4	75.1	4679	13	AF043303

ALIGNMENTS

RESULT 1
BD242766
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD242766 4718 bp DNA linear PAT 17-JUL-2003
Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same.
BD242766
BD242766.1 GI:33052536
JP 2002529098-A/1.
unidentified
unidentified
unclassified.
1 (bases 1 to 4718)
Wilson, J.M. and Xiao, W.
Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same
Patent: JP 2002529098-A 1 10-SEP-2002.
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
OS AAV-1
PN JP 2002529098-A/1
PD 10-SEP-2002
PP 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES M WILSON, WEIDONG XIAO
PC C12N15/09, A61K31/711, A61P43/00, A61P43/00, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02//A61K35/76, C12N15/00, C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
cell containing the same
Location/Qualifiers
FH Key (335) . (2206)
FT CDS (2223) . (4430) .
FT CDS Location/Qualifiers
1 . 4718
/organism="unidentified"
/mol_type="Genomic DNA"
/db_xref="taxon:32644"

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;

```
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAAAGTCCGC 60
Db 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAAAGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143

RESULT 2
AR562498
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same
JOURNAL Patent: US 6759237-A 1 06-JUL-2004;
The Trustees of The University of Pennsylvania; Philadelphia, PA
FEATURES
source
1. .4718
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAAAGTCCGC 60
Db 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAAAGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143

RESULT 3
AR753251
LOCUS AR753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
ACCESSION AR753251
VERSION AR753251.1 GI:32166108
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1
AUTHORS Gao,G., Wilson,J.M. and Alvira,M.
TITLE A method of detecting and/or identifying adeno-associated virus (AAV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
source
1. .4718
/organism="Adeno-associated virus 1"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:85106"

ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAAAGTCCGC 60
Db 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAAAGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143

RESULT 4
AF063497
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999
DEFINITION Adeno-associated virus 1, complete genome.
ACCESSION AF063497
VERSION AF063497.1 GI:4689096
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1 (bases 1 to 4718)
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and Wilson,J.M.
TITLE Gene therapy vectors based on adeno-associated virus type 1
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)
PUBMED 10196295
REFERENCE 2 (bases 1 to 4718)
AUTHORS Xiao,W. and Wilson,J.M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA
FEATURES
source
1. .4718
/organism="Adeno-associated virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:85106"
/note="AAV1"
335..2206
/codon_start=1
/product="nonstructural protein"
/protein_id="AAD27758.1"
/db_xref="GI:4689098"
/translation="MPGYEIVIKVPSDLDEHLPGISDFVSWAEKEWELPPDSMD
LNLHQAPLTAELKORDFLVQMRVSKAPALRVQPEKGSYFHLHLVETTVGVS
MYLGRFLSQIRDLKLVQTIYRGIEPLFPNFAVTKTRNGAGGKNVDSYIPNYLLPK
TQPELQWATMMSEYISACLNLAERKLVLAHQRLTHVTSQONKENLNPNSDAPVRS
KTSAPYLVGPAPPADIKTNRIYRIELNGYEVAGSVFLGWAQKRFKRTIWLPG
PATGTGNTIAEIAHAAVFPVGCVMNTNFPCDVKMVIWMSEGMTAKVBSAKA
ILGSKVRVDQKCKSSAOIDPTPIVTNTNMCAVIDGNSTTFEHQOPLQDRMKFEL
TRLEHDFGKVTQKEVFFRFAQDHVTEVAHEFVVRKGANKRPAPDDADSEKRA
CPSVAPSTSDAEGAPVDADRIYQNKSRHAGMLQMLFPCKTCERMNQNFICFTHGT
RDCSCFFGVSESQPVVRKRTYRKLCAIHLHLLGRAPEIACSAACDLVNVLDLDCVSEQ"
2223..4433
/codon_start=1
/product="capsid protein"
/protein_id="AAD27757.1"
/db_xref="GI:4689097"
/translation="MAADGYLPDWLEDLNLSGIREWMDLKPQAPKPKANQKQDDGRG
LVLPGYKYLGPFGNLDKGEVPAADAAALSHDKAYDQQLKAGDNPYLRYNHADAFOE
RLQEDTSFGNLGRAVFOAKRVLEPLGLVBEAGAKTAPGKKRPVQSPOEPDSSSGIG
KTGQOPAKRLNFGQTGDSSEVPDPQLGEPPTAAVGPPTTMSAGGGAPMADNNEGA
```

DGVGNAGNWHCDSTWLGDRVITTRTWALPTYNHLYKQISSASTGASNDNHFGY
STPMGYFDNFRPHCFSPRDMORLNNNGWFRPKLNFKNIOVKQEVTTNDGVTTIA
NNLSTVQVFSDEYQLPYVLGSAHQCLPPFADVFMIPQYGYLTNNNGSAVGRSS
FYCLEYFSPQMLRTGNFTFSYTFEEVPHSHSYAHSQSIDRLMNPILDOYLYLARTQ
NOSGSAQNDLFLSRGSPAGMSVQPKNWLPGCYRQQRVSKTKTNNNSNFTWTGASK
YNLNGRESIINPGTMAWASHKDDKFFPMGSMVIFGKESAGASNTALDNVMTDDEEI
TADHPVATFRFTGTVANFQSSSTDPATGDVHMALPGMWQDRDYLQGPWAKIPIH
TQGHFHPSPLMGFGFLKPNPPQILIKNTVPANPPAEFSATKFAFPIQYSTGVQSV
IEWELQKENSXKWNPEVQYTSYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL"

ORIGIN

Query Match 100.0%; Score 143; DB 13; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCCCTGCGGACCAAGGTCCGC 60

Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCCCTGCGGACCAAGGTCCGC 60

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGAGCGCGAGAGGAGTG 120

Db 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGAGCGAGAGGAGTG 120

QY 121 GGCAACTCCATCTAGGGGTAA 143

Db 121 GGCAACTCCATCTAGGGGTAA 143

RESULT 5
AF028704/c

LOCUS Adeno-associated virus 6, complete genome. 4683 bp DNA linear VRL 12-JAN-1998

DEFINITION Adeno-associated virus 6, complete genome.

ACCESSION AF028704

VERSION AF028704.1 GI:2766605

KEYWORDS Adeno-associated virus 6

SOURCE Adeno-associated virus 6

ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

REFERENCE Rutledge, E.A., Halbert, C.L. and Russell, D.W.

AUTHORS Infectious clones and vectors derived from adeno-associated virus

TITLE (AAV) serotypes other than AAV type 2

J. Virol. 72 (1), 309-319 (1998)

PUBLISHED 9420229

REFERENCE 2 (bases 1 to 4683)

AUTHORS Rutledge, E.A. and Russell, D.W.

TITLE Direct Submission

JOURNAL Submitted (24-SEP-1997) Hematology, Univ. of Washington, Box

357720, Seattle, WA 98195, USA

FEATURES Location/Qualifiers

1..4683

/organism="Adeno-associated virus 6"

/mol_type="genomic DNA"

/db_xref="taxon:68558"

/note="AAV6"

320..2191

/codon_start=1

/evidence=not_experimental

/product="nonstructural protein Rep78"

/protein_id="AAB95449.1"

/db_xref="GI:2766606"

/translations="MPGFYEVIVKPSDLDLHLPGLISDFVNVVAEKWELPPDSMD

LNLEQAPLTVAEKLQDFLQVRRYSKAPALFFVQFEGESYFHLHLIVETGVKS

MVLGRFLSRQLQVLTITYGIEPTLPNFAVTKTRNGAGGNKVDCEVYNLLPK

TPQLQWAMTMEYISACLNLAERKLVAHDLTHVSQTOEKENLNPNSDAPVIRS

KTSARYMELVQVITSEKQWIOEDQASYISFNAASNSRSQIKAAIDNAGKIMAL

ORIGIN

Query Match 97.1%; Score 138.8; DB 13; Length 4683;

Best Local Similarity 98.6%; Pred. No. 2e-20;

Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCCCTGCGGACCAAGGTCCGC 60

Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCCCTGCGGACCAAGGTCCGC 4624

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGAGCGAGAGGAGTG 120

Db 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGAGCGAGGAGTG 4564

QY 121 GGCAACTCCATCTAGGGGTAA 142

Db 4563 GGCAACTCCATCTAGGGGTAA 4542

RESULT 6
BD242775/c

LOCUS Adeno-associated virus serum type 1 nucleic acid sequence, vector

DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector

ACCESSION BD242775

VERSION BD242775.1 GI:33052545

KEYWORDS JP 2002529098-A/10.

SOURCE unidentified

ORGANISM unidentified

REFERENCE 1 (bases 1 to 4683)

AUTHORS Wilson, J.M. and Xiao, W.

TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector

JOURNAL Patent: JP 2002529098-A 10 10-SEP-2002;

COMMENT THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA

OS aav-6

PN JP 2002529098-A/10

PD 10-SEP-2002

PP 02-NOV-1999 JP 2000581227

PR 05-NOV-1998 US 60/107114

PI JAMES M WILSON, WEIDONG XIAO

PC C12N15/09, A61K31/711, A61K48/00, A61P43/00, C12N1/15, C12N1/19, PC

C12N1/21,

PC C12N5/10, C12P21/02//A61K35/76, C12N15/00, C12N5/00 CC

Adeno-associated virus serum type 1 nucleic acid sequence, CC

vector and host

CC cell containing the same

FH Key Location/Qualifiers

FT source 1..4683

FT /organism='aav-6'.

Location/Qualifiers

1..4683

/organism="unidentified"

/mol_type="genomic DNA"

2208..4418

CDS

FEATURES

source

1..4683

/organism="unidentified"

/mol_type="genomic DNA"

2208..4418

CDS

FEATURES

source


```
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH Key Location/Qualifiers
FT CDS (335)..(2206)
FT CDS (2223)..(4430).
FEATURES
    source
        1..4718
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 3.7e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCGCG 60
Db 4718 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCGCG 4659
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 4658 CGTCTCGGACCTTTGGTCCGAGGCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599
QY 121 GGCACTCCATCACTAGGGGTAA 143
Db 4598 GGCACTCCATCACTAGGGGTAA 4576
RESULT 13
AR562498/C
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
JOURNAL Patent: US 6759237-A 1 08-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
    source
        1..4718
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 3.7e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCGCG 60
Db 4718 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCGCG 4659
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 4658 CGTCTCGGACCTTTGGTCCGAGGCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599
QY 121 GGCACTCCATCACTAGGGGTAA 143
Db 4598 GGCACTCCATCACTAGGGGTAA 4576
RESULT 14
AX753251/C
LOCUS AX753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
```

```
ACCESSION AX753251
VERSION AX753251.1 GI:32166108
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1
AUTHORS Gao,G., Wilson,J.M. and Alvira,M.
TITLE A method of detecting and/or identifying adeno-associated virus
(AVV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
    Location/Qualifiers
        source
            1..4718
                /organism="Adeno-associated virus 1"
                /mol_type="unassigned DNA"
                /db_xref="taxon:85106"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 3.7e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCGCG 60
Db 4718 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCGCG 4659
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 4658 CGTCTCGGACCTTTGGTCCGAGGCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599
QY 121 GGCACTCCATCACTAGGGGTAA 143
Db 4598 GGCACTCCATCACTAGGGGTAA 4576
RESULT 15
AF063497/C
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999
DEFINITION Adeno-associated virus 1, complete genome.
ACCESSION AF063497
VERSION AF063497.1 GI:4689096
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1 (bases 1 to 4718)
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and
Wilson,J.M.
TITLE Gene therapy vectors based on adeno-associated virus type 1
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)
PUBMED 10196295
REFERENCE 2 (bases 1 to 4718)
AUTHORS Xiao,W. and Wilson,J.M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601
Spruce Street, Philadelphia, PA 19104, USA
FEATURES
    Location/Qualifiers
        1..4718
            /organism="Adeno-associated virus 1"
            /mol_type="genomic DNA"
            /db_xref="taxon:85106"
            /note="AAV1"
            335..2206
                /codon_start=1
                /product="nonstructural protein"
                /protein_id="AAD27758.1"
                /db_xref="GI:4689098"
                /translation="MPGYEIVIKVPSDLDEHLPGISDFVSWAEKEWELPPDSMD
                LNLIEQAPLTVAELKLDPLVOMRRVSKAPEALFFVQFEKESYFHLHLLVSTTVK
                MYLGRFLSLQIRDKLVQTTIYRGTEPLPNMFAVTKTRNGAGGKGVDBCYIYNYLLPK
                TOPELQAWTMMEEYISACLNLAEKRLVAOHLTHVSQTQEQNKENLNPNSDAPVIRS
                KTSARYMELVGLVDRGITSEKQWIQEDQASVYSFNAASNSRSQIKALDINAGKIMAL"
CDS
```

TKSADYLVGPAPPADIKTNRIYRIELENGYEPAYAGSVFLGWAQKRFGRNTIWLFG
PATTKTNIADIAIAHAFYPCVNTNENFPENDCVDKMVIWEEGRMTAKRVESAKA
ILGSKVRDQKCKSAQIDPTPTVTNTNCAVIDGNSTTFEHQDQLQDRMFPEL
TRRLBHDCKVTKQVKEPFRWAQDHTEVAHEFYVRKGANKRPAPDDADKSEPKRA
CPSVADPSTDAEGAPVDPADRYONKCSRHAGMLOMLPCKTCERNONFNICFTHGT
RDCSECFQGVSESQPVVRKRYRKLCIAIHLLGRAPEIACSDLVNVDLDDCVSEQ"
2223.. .4433
/codon_start=1
/product="capsid protein"
/protein_id="AAD27757.1"
/db_xref="GI:4689097"
/translation="MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRG
LVLPGYKILGPPNGLDKGEFVNAADAAALHDKAYDQLKAGDNPYLRYNHADAEFQE
RLQEDTSFGNIGRAVFOAKRVLEPLGLVEEGAKTAPGKRKPVQSQEPDSSSIG
KTQOPAKKRLNFGOTGSESVDPQPLCEPPATPAAGVPTTMASSGGA PMADNNEGA
DGVNAGNWHCDSTWLGDRVITSTRTWALPTNNHLYKQISSASTGASNDNHVFGY
STPMGYFDNRRCHPSPRDQRLNNWGFEPKRLNKLNIQVKEVTNDGVTTIA
NNLTSTVQVFSSETQLPYVLGSAHQCLPPFPADVPMIPQYGYLTLANGSAVGRSS
FYCLEYFSPQMLRTGNFTSYTFEEVPHSHSQSLDRMLNPLIDQYLYLNRTQ
NQSGAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQVRVSKTKTDNNNSNFTWTGASK
YNLNGRESTINPGTAMASHKDDKFFPMSGVMIIFGESAGASNTALDNVMTDBEEI
KATNPVATERFCTVAVNPQSSSTDPATGDVHAMGALPGMWODRDVYLQGPWAKIPH
TDGHFHPSLMGFGGLKNPPQIILKNTFVPANPPAEFSATKFASTFYQYSTGQVSVE
IEWELQKENSKNWNPVQYTSNYSKANSVDFTVDNNGLYTEPRPIGTRYLTRPL"

ORIGIN

Query Match 83.2%; Score 119; DB 13; Length 4718;
Best Local Similarity 89.5%; Pred. No. 3.7e-16;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGCCCCACTCCTCTCTGCGCGCTCGCTCGGTGGGGCCTGCGGACCAAAGGTCCGC 60
DB 4718 TTGCCCCACTCCTCTCTGCGCGCTCGCTCGGTGGGGCCTGCGGACCAAAGGTCTGCG 4659
QY 61 AGACGGCAGAGCTCTCTCTGCGGCGCCACCGAGCGAGCGCGGAGAGGAGTG 120
DB 4658 CGTCTCGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGCGGAGGAGTG 4599
QY 121 GGCACTCCATCACTAGGGGTAA 143
DB 4598 GGCACTCCATCACTAGGGGTAA 4576

Search completed: November 29, 2005, 00:05:25
Job time : 524.566 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 08:46:14 ; Search time 66.9124 Seconds
(without alignments)
14243.261 Million cell updates/sec

Title: US-10-696-900-1_COPY_1_143

Perfect score: 143

Sequence: 1 ttgccaccctcctctctcg.....aactccatcactagggttaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	143	3	AAD00781
2	143	100.0	4718	3	AAD00772
3	143	100.0	4718	10	AD876507
4	143	100.0	4718	10	ADL13984
5	143	100.0	4718	12	ADG39758
6	138.8	97.1	4683	4	AAF23749
7	138.8	97.1	4683	10	ADL13983
8	138.8	97.1	4683	12	ADG39763
9	135.6	94.8	4721	10	ADG76502
10	135.6	94.8	4721	12	ADG39764
11	135.6	94.8	4721	14	ADV67509
12	135.6	94.8	4721	14	ADZ27030
13	119	83.2	4718	3	AAD00772
14	119	83.2	4718	10	AD876507
15	119	83.2	4718	10	ADL13984
16	119	83.2	4718	12	ADG39758
17	113.2	79.2	4722	4	AAF23748
18	113.2	79.2	4722	12	ADG39760
19	113.2	79.2	4726	10	AD876509

20	113.2	79.2	4726	10	ADL13985
21	113.2	79.2	4726	12	ADG39759
22	112.2	78.5	4722	4	AAF23748
23	112.2	78.5	4722	12	ADG39760
24	111.6	78.0	4721	10	AD876502
25	111.6	78.0	4721	12	ADG39764
26	111.6	78.0	4721	14	ADV67509
27	111.6	78.0	4721	14	ADZ27030
28	107.4	75.1	145	2	AAQ41448
29	107.4	75.1	145	2	AAT03385
30	107.4	75.1	145	6	ABX34295
31	107.4	75.1	145	6	ABS69884
32	107.4	75.1	145	8	ACC58491
33	107.4	75.1	145	9	ACF35876
34	107.4	75.1	145	13	ADM93250
35	107.4	75.1	146	8	ABX93568
36	107.4	75.1	165	2	AAQ66769
37	107.4	75.1	165	2	AAT49462
38	107.4	75.1	165	4	AAD03535
39	107.4	75.1	165	6	ABS69886
40	107.4	75.1	165	8	ABV77279
41	107.4	75.1	170	12	ADL91189
42	107.4	75.1	207	6	AA44621
43	107.4	75.1	272	3	AZ47166
44	107.4	75.1	272	4	AAD03534
45	107.4	75.1	955	6	AAD37254

ALIGNMENTS

RESULT 1

AAD00781

ID AAD00781 standard; DNA; 143 BP.

XX AAD00781;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 5' inverted terminal repeat.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;

KW cap protein; recombinant viral vector; Gene delivery; gene therapy;

KW vaccine; transgene; inverted terminal repeat; ITR; ss.

XX Adeno associated virus serotype 1.

XX Key Location/Qualifiers

FT misc_binding 1..41

FT /tag= a

FT /bound_moiety= "binds to nucleotides 85..125"

FT stem_loop 42..62

FT /tag= b

FT stem_loop 64..84

FT /tag= c

FT misc_binding 85..125

FT /tag= d

FT /bound_moiety= "binds to nucleotides 41..1"

XX WO200028061-A2.

XX 18-MAY-2000.

XX 02-NOV-1999; 99WO-US025694.

XX 05-NOV-1998; 98US-0107114P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Xiao W;

XX WPI; 2000-376571/32.

XX

```
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host.
XX
PS Example 2; Fig 2; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antisense delivery vectors. The AAV-1 does not induce the
CC formation of neutralising antibodies specific to any serotype of AAV
CC hence is useful for transforming host cells, and in the preparation of a
CC medicament for the delivery of transgene to a host. The present sequence
CC is a 5' ITR of AAV-1 DNA which is useful in the production of recombinant
CC viral vector. The ITR forms a T-shaped hairpin structure
XX
SQ Sequence 143 BP; 24 A; 50 C; 48 G; 21 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 3; Length 143;
Best Local Similarity 100.0%; Pred. No. 2,3e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCACCTCCCTCTCTGCGCGCTCGCTCGCTGCGTGGGGCTTGGGACCAAGGTCGCC 60
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 TTGCCACCTCCCTCTCTGCGCGCTCGCTCGCTGCGTGGGGCTTGGGACCAAGGTCGCC 60
QY 61 AGACGCGAGAGCTCTCTGCGCGCCCGCCACCGAGCGGCGCGCAGAGGAGTG 120
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 AGACGCGAGAGCTCTCTGCGCGCCCGCCACCGAGCGGCGCGCAGAGGAGTG 120
QY 121 GGCACCTCCATCACTAGGGGTAA 143
Db | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 GGCACCTCCATCACTAGGGGTAA 143
RESULT 2
ID AAD00772 standard; DNA; 4718 BP.
XX
AC AAD00772;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 DNA.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; ss.
XX
OS Adeno associated virus serotype 1.
XX
FH Key Location/Qualifiers
FT repeat_unit 1..143
FT /tag= a
FT /label= 5' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT protein_bind 89..110
FT /tag= b
FT /bound moiety= "Rep protein"
FT misc_feature 124..125
FT /tag= c
FT /note= "Terminal resolute site (TRS)"
FT misc_binding 219..226
FT /tag= d
FT /bound moiety= "USF"
FT promoter 236..299
FT /tag= e
```

```
FT /label= P5_promoter
FT 237..245
FT /tag= f
FT /bound moiety= "YY1 factor"
FT 270..275
FT /tag= g
FT /label= P5_TATA-Box
FT 299..306
FT /tag= h
FT /note= "YY1/p5 RNA"
FT 335..2272
FT /tag= i
FT /product= "Rep 68"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT 335..2206
FT /tag= j
FT /product= "Rep 78"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT 857..862
FT /tag= k
FT /label= P19_TATA_Box
FT 882..883
FT /tag= m
FT /note= "P19 RNA"
FT 1007..2272
FT /tag= o
FT /product= "Rep 40"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT 1007..2206
FT /tag= n
FT /product= "Rep 52"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT 1836..1841
FT /tag= p
FT /label= P40_TATA-BOX
FT 1875..1876
FT /tag= q
FT /note= "P40 RNA"
FT 1924..2220
FT /tag= k
FT /note= "This region interrupts the coding sequence of Rep
FT 68 and Rep 40"
FT 2223..4433
FT /tag= r
FT /product= "VP1 protein"
FT /note= "Capsid protein"
FT 2634..4433
FT /tag= b
FT /product= "VP2 protein"
FT /note= "Capsid protein"
FT /partial
FT 2829..4433
FT /tag= t
FT /product= "VP3 protein"
FT /note= "Capsid protein"
FT 4447..4452
FT /tag= u
FT 4576..4718
FT /tag= v
FT /label= 3' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
XX
XX WO200028061-A2.
XX 18-MAY-2000.
XX
```

PF 02-NOV-1999; 99WO-US025694.
XX
PR 05-NOV-1998; 98US-0107114P.
PR
PA (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Xiao W;
PI
XX WPI: 2000-376571/32.
DR
DR P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
DR AAY71169.
XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host.
XX
XX Claim 1; Fig 1; 108pp; English.
XX
XX The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
CC characterised by two inverted terminal repeats (ITR) and open reading
CC frames for rep and capsid (cap) proteins. The rep reading frame encodes
CC four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
CC frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
CC sequence or its fragments particularly ITRs, rep and cap coding regions,
CC are useful in production of recombinant viral vectors for gene delivery.
CC These vectors can be used as gene therapy vectors, vaccine vectors or
CC antisense delivery vectors. The AAV-1 does not induce the formation of
CC neutralising antibodies specific to any serotype of AAV hence is useful
CC for transforming host cells, and in the preparation of a medicament for
CC the delivery of transgene to a host
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 3; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.8e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCGCTCGCGACCAAAAGGTCCGC 60
Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCGCTCGCGACCAAAAGGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143
RESULT 3
ADE76507
ID ADE76507 standard; DNA; 4718 BP.
XX
AC ADE76507;
XX
DT 29-JAN-2004 (first entry)
XX
XX Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
DE
XX adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antiinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
OS Adeno-associated virus 1.
XX
XX EF1310571-A2.
XX
XX 14-MAY-2003.
XX
XX 12-NOV-2002; 2002EP-00257826.
PF

XX 13-NOV-2001; 2001US-0350607P.
PR 17-DEC-2001; 2001US-0341117P.
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386675P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
PA
XX Gao G, Wilson JM, Alvira M;
XX WPI: 2003-450984/43.
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX Claim 14; SEQ ID NO 6; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antirheumatic, antithyroid, dermatological, and antinflammatory. The AAV
CC antidiabetic, antithyroid, dermatological, and antinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.8e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCGCTCGCGACCAAAAGGTCCGC 60
Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCGCTCGCGACCAAAAGGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143
RESULT 4
ADL13984
ID ADL13984 standard; DNA; 4718 BP.
XX
XX ADL13984;
XX
XX 06-MAY-2004 (first entry)
XX
XX Adeno-associated virus serotype 1 complete DNA.
DE
XX ss; cytostatic; neuroprotective; antinflammatory; gene therapy;
KW expression construct; adeno-associated virus;
KW integration efficiency element; inverted terminal repeat; integration;
KW chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;
KW retinoblastoma; inflammatory disease; arthritis;
KW neurodegenerative disease.
XX
XX Adeno-associated virus 1.
OS
XX WO2003087334-A2.
PN

XX PD 23-OCT-2003.
XX PF 09-APR-2003; 2003WO-US011191.
XX PR 09-APR-2002; 2002US-0371044P.
XX PA (CORR) CORNELL RES FOUND INC.
XX PI Falck-Pedersen ES, Philipott N;
XX DR WPI; 2003-833723/77.
XX PT New expression construct comprising a nucleic acid sequence encoding an
PT adeno-associated virus integration efficiency element, useful for
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease
PT e.g. arthritis.
XX PS Disclosure; SEQ ID NO 3; 62pp; English.
XX CC The invention relates to an expression construct comprising a nucleic
CC acid sequence encoding an adeno-associated virus integration efficiency
CC element (AAV IBE), which is devoid of AAV inverted terminal repeats (AAV
CC ITRs) and site-specifically integrates into a host cell chromosome when
CC provided to the host cell in conjunction with an AAV Rep protein. The
CC expression construct can be used as a therapeutic factor for treating a
CC mammal for a pathologic state which is cancer, including lung cancer,
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,
CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic
CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic
CC state includes inflammatory disease (arthritis), neurodegenerative
CC disease, a disease of an organ attributed to the presence of increased or
CC decreased level of a particular gene product(s). This sequence
CC corresponds to the AAV serotype 1 complete DNA including the IEE
CC sequence.
XX SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.8e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 60
Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGGCGAGAGGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGGCGAGAGGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143
RESULT 5
ADG39758
ID ADG39758 standard; DNA; 4718 BP.
XX AC ADG39758;
XX DT 11-MAR-2004 (first entry)
XX DE AAV-1 genomic DNA sequence SEQ ID NO:20.
XX KW parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
KW recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.

XX OS Adeno-associated virus 1.
XX PN WO2003104392-A2.
XX PD 18-DEC-2003.
XX PF 02-DEC-2002; 2002WO-US038423.
XX PR 18-DEC-2001; 2001US-0341919P.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PI Samulski RJ, Rabinowitz JE;
XX DR WPI; 2004-062324/06.
XX PT New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX PS Disclosure; SEQ ID NO 20; 115pp; English.
XX CC The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (I); (2) a cell comprising (I), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprises sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 12; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.8e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 60
Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGGCGAGAGGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGGCGAGAGGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143
RESULT 6
AAF23749/c
ID AAF23749 standard; DNA; 4683 BP.
XX AC AAF23749;
XX AC AAF23749;

XX 28-MAR-2001 (first entry)
XX AAV6 DNA sequence.
XX AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anemia; thalassemia;
KW blood clotting disorder; diabetes; ss.
XX Adeno associated virus.
XX US6156303-A.
XX 05-DEC-2000.
XX 11-JUN-1997; 97US-00873168.
XX 11-JUN-1997; 97US-00873168.
XX (UNIW) UNIV WASHINGTON.
XX Russell DW, Rutledge EA;
XX WPI; 2001-060164/07.
XX Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes.
XX Claim 1; Fig 1; 50pp; English.
XX The present invention relates to adeno-associated virus serotypes. The
XX present sequence is the DNA sequence of one such serotype (AAV6). AAV6
XX can be used to construct AAV viral vectors for use in gene therapy for a
XX range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
XX sickle cell anemia, thalassemia, blood clotting disorders and diabetes.
XX The AAV viral vectors have increased transduction efficiency of a
XX particular host cell as the AAV virion containing the AAV vector genome
XX can be modified to express a capsid protein of an AAV serotype that
XX transduces the selected host cell
XX
XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
Query Match 97.1%; Score 138.8; DB 4; Length 4683;
Best Local Similarity 98.6%; Pred. No. 3.5e-27;
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTGCCCACTCCCTCTGCGCGCTCGCTCGCTGGGGCGCTCGCGACCAAGGTCGCG 60
Db 4683 TTGCCCACTCCCTCTATGCGCGCTCGCTCGCTGGGGCGCTCGCGACCAAGGTCGCG 4624
QY 61 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 120
Db 4623 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 4564
QY 121 GGCAACTCCATCACTAGGGTA 142
Db 4563 GGCAACTCCATCACTAGGGTA 4542
RESULT 7
ADL13983/c
ID ADL13983 standard; DNA; 4683 BP.
XX ADL13983;
AC
DT 06-MAY-2004 (first entry)
XX Adeno-associated virus serotype 6 complete DNA.
XX ss; cytosatic; neuroprotective; antiinflammatory; gene therapy;
KW expression construct; adeno-associated virus;
KW integration efficiency element; inverted terminal repeat; integration;

KW Chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;
KW retinoblastoma; inflammatory disease; arthritis;
KW neurodegenerative disease.
XX Adeno-associated virus 6.
XX WO2003087334-A2.
XX 23-OCT-2003.
XX 09-APR-2003; 2003WO-US011191.
XX 09-APR-2002; 2002US-0371044P.
XX (CORR) CORNELL RES FOUND INC.
XX Falck-Pedersen ES, Philpott N;
XX WPI; 2003-833723/77.
XX New expression construct comprising a nucleic acid sequence encoding an
PT adeno-associated virus integration efficiency element, useful for
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease
PT e.g. arthritis.
XX Disclosure; SEQ ID NO 2; 62pp; English.
XX The invention relates to an expression construct comprising a nucleic
XX acid sequence encoding an adeno-associated virus integration efficiency
XX element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV
XX ITRs) and site-specifically integrates into a host cell chromosome when
XX provided to the host cell in conjunction with an AAV Rep protein. The
XX expression construct can be used as a therapeutic factor for treating a
XX mammal for a pathologic state which is cancer, including lung cancer,
XX colon cancer, renal cancer, anal cancer, bile duct cancer, bladder
XX cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,
XX cervical cancer, lymphoma, endometrial cancer, esophageal cancer,
XX gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,
XX liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic
XX cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,
XX testicular cancer, thymus cancer or thyroid cancer. Other pathologic
XX state includes inflammatory disease (arthritis), neurodegenerative
XX disease, a disease of an organ attributed to the presence of increased or
XX decreased level of a particular gene product(s). This sequence
XX corresponds to the AAV serotype 6 complete DNA including the IEE
XX sequence.
XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
Query Match 97.1%; Score 138.8; DB 10; Length 4683;
Best Local Similarity 98.6%; Pred. No. 3.5e-27;
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTGCCCACTCCCTCTGCGCGCTCGCTCGCTGGGGCGCTCGCGACCAAGGTCGCG 60
Db 4683 TTGCCCACTCCCTCTATGCGCGCTCGCTCGCTGGGGCGCTCGCGACCAAGGTCGCG 4624
QY 61 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 120
Db 4623 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 4564
QY 121 GGCAACTCCATCACTAGGGTA 142
Db 4563 GGCAACTCCATCACTAGGGTA 4542
RESULT 8
ADG39763/c
ID ADG39763 standard; DNA; 4683 BP.
XX ADG39763;
AC
XX 11-MAR-2004 (first entry)
DT

```
XX AAV-6 genomic DNA sequence SEQ ID NO:25.
DE parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
XX recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
OS Adeno-associated virus 6.
XX
PN WO2003104392-A2.
XX
PD 18-DEC-2003.
XX
XX 02-DEC-2002; 2002WO-08038423.
XX
XX 18-DEC-2001; 2001US-0341919P.
XX
XX (UYNC-) UNIV NORTH CAROLINA.
XX
XX Samulski RJ, Rabinowitz JE;
XX
XX WPI; 2004-062324/06.
XX
XX New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
XX Disclosure; SEQ ID NO 25; 115pp; English.
XX
XX The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
SQ
Query Match 97.1%; Score 138.8; DB 12; Length 4683;
Best Local Similarity 98.6%; Pred. No. 3.5e-27;
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 60
Db 4683 TTGCCACTCCCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 4624
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCCACCCAGCGAGCGGCGAGAGGAGTG 120
Db 4623 AGACGGCAGAGCTCTGCTTGGCGGCCACCCAGCGAGCGGCGATAGAGGAGTG 4564
QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 4563 GGCAACTCCATCACTAGGGGTA 4542
```

```
RESULT 9
ADE76502
ID ADE76502 standard; DNA; 4721 BP.
XX
XX ADE76502;
AC
XX 29-JAN-2004 (first entry)
DT
XX
XX Adeno-associated virus (AAV) related DNA, SEQ ID No 1.
DE
XX
XX adeno-associated virus; AAV; cytotatic; antipsoriatic; antirheumatic;
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
XX Adeno-associated virus 7.
OS
XX
XX EP1310571-A2.
PN
XX
XX 14-MAY-2003.
PD
XX
XX 12-NOV-2002; 2002EP-00257826.
PF
XX
XX 13-NOV-2001; 2001US-0350607P.
PR
XX 17-DEC-2001; 2001US-0341117P.
PR
XX 01-MAY-2002; 2002US-0377066P.
PR
XX 05-JUN-2002; 2002US-0386675P.
PR
XX
XX (UYPE-) UNIV PENNSYLVANIA.
PA
XX
XX Gao G, Wilson JM, Alvira M;
PI
XX
XX WPI; 2003-450984/43.
P-PSDB; ADE76503.
DR
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX Claim 17; SEQ ID NO 1; 419pp; English.
PS
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytotatic,
CC antipsoriatic, antirheumatic, antithyroid, and antinflammatory. The AAV
CC antidiabetic, antithyroid, dermatological, and antinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
XX Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;
SQ
Query Match 94.8%; Score 135.6; DB 10; Length 4721;
Best Local Similarity 97.2%; Pred. No. 2.5e-26;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 60
Db 1 TTGCCACTCCCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCCACCCAGCGAGCGGCGAGAGGAGTG 120
```

Db 61 AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGGCGAGCGGCGCATAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 121 GGCAACTCCATCACTAGGGGTA 142
RESULT 10
ADG39764
ID ADG39764 standard; DNA; 4721 BP.
XX
AC ADG39764;
XX
DT 11-MAR-2004 (first entry)
XX
DE AAV-7 genomic DNA sequence SEQ ID NO:26.
XX
KW parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
KW recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
OS Adeno-associated virus 7.
XX
PN WO2003104392-A2.
XX
PD 18-DEC-2003.
XX
PF 02-DEC-2002; 2002WO-US038423.
XX
PR 18-DEC-2001; 2001US-0341919P.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Samulski RJ, Rabinowitz JE;
XX
DR WPI; 2004-062324/06.
XX
XX New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
PS Disclosure; SEQ ID NO 26; 115pp; English.
XX
CC The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (1) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;
Query Match 94.8%; Score 135.6; DB 12; Length 4721;

Best Local Similarity 97.2%; Pred. No. 2.5e-26;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 TTGCGCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCTGCGGACCAAGGTCCGC 60
Db 1 TTGCGCACTCCCTCTCTATGCGCGCTCGCTCGCTCGGTGGGGCCTGCGGACCAAGGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGAGCGAGCGCGCATAGAGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGAGCGAGCGCGCATAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 121 GGCAACTCCATCACTAGGGGTA 142
RESULT 11
ADV67509
ID ADV67509 standard; DNA; 4721 BP.
XX
AC ADV67509;
XX
DT 10-MAR-2005 (first entry)
XX
DE Nucleotide sequence of AAV serotype 7.
XX
KW antiarteriosclerotic; antilipemic; gene therapy; cholesterol;
KW apolipoprotein E; apoE; apolipoprotein A; apoA; atherosclerosis;
KW lipoprotein defect; ss.
XX
OS Adeno-associated virus.
XX
FH Key Location/Qualifiers
FT repeat_region 1..107
FT /*tag= a, ITR"
FT /*note= "5', 3049
FT misc_feature 825..3049
FT /*tag= b
FT /*note= "encodes vp1"
FT misc_feature 1234..3049
FT /*tag= c
FT /*note= "encodes vp2"
FT misc_feature 1434..3049
FT /*tag= d
FT /*note= "encodes vp3"
FT repeat_region 4704..4721
FT /*tag= e
FT /*note= "3', ITR"
XX
WO2004108922-A2.
XX
PD 16-DEC-2004.
XX
PF 23-APR-2004; 2004WO-US010965.
XX
PR 25-APR-2003; 2003US-0465293P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Rader DJ, Wilson JM;
XX
DR WPI; 2005-031700/03.
XX
PT Lowering total cholesterol levels and treating atherosclerosis in a
PT subject comprises delivering a recombinant adeno-associated virus (AAV)
PT comprising an AAV serotype capsid protein or a gene encoding human
PT apolipoprotein E (apoE) or apoA.
XX
PS Disclosure; SEQ ID NO 7; 69pp; English.
XX
CC The specification describes a method for lowering total cholesterol
CC levels in a subject. The method comprises delivering to the subject a
CC recombinant adeno-associated virus (AAV) comprising a gene encoding a

CC human apolipoprotein B (apoB) or apoA under the control of a regulatory
 CC control sequence which directs expression of the gene. The recombinant
 CC AAV also comprises a capsid protein selected from an AAV serotype (e.g.
 CC AAV7 or AAV8) which preferentially expresses high levels of transgene in
 CC live. A therapeutically effective amount of apoE or apoA expression is
 CC obtained upon delivery of low dose of AAV. The method of the invention is
 CC useful for lowering total cholesterol levels in a subject, e.g. for
 CC treating atherosclerosis, and for correcting defects in lipoprotein. The
 CC present sequence represents AAV serotype 7 (AAV7), which may be used to
 CC produce recombinant AAV vectors of the invention.

XX
 SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;

Query Match 94.8%; Score 135.6; DB 14; Length 4721;
 Best Local Similarity 97.2%; Pred. No. 2.5e-26;
 Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TTGCCACCTCCCTCTCTGCGGCTCGCTCGGTCGGTGGGGCTGCGGACCAAGGTCCGC 60
 DB 1 TTGGCCACTCCCTCTCTATGCGGCTCGCTCGGTCGGTGGGGCTGCGGACCAAGGTCCGC 60
 QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCAGAGAGGAGTG 120
 DB 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCATAGAGGAGTG 120
 QY 121 GGCACCTCCATCACTAGGGGTA 142
 DB 121 GCCAACTCCATCACTAGGGGTA 142

RESULT 12

AD227030
 ID AD227030 standard; DNA; 4721 BP.

XX
 AC AD227030;

XX 30-JUN-2005 (first entry)

XX Adeno-associated virus DNA SEQ ID NO 180.

XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
 KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
 KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
 KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
 KW virucide; antibacterial; cytostatic; antitumor; dermatological; ds.

XX Adeno-associated virus.

XX WO200503321-A2.

XX 14-APR-2005.

XX 30-SEP-2004; 2004WO-US028817.

XX 30-SEP-2003; 2003US-0508226P.

XX 29-APR-2004; 2004US-0566546P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX WPI; 2005-285437/29.

XX New adeno-associated virus (AAV) clade comprising at least three AAV
 PT members, useful for preventing and/or treating arthritis, multiple
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
 PT infection and cancer.

XX Claim 12; SEQ ID NO 180; 569pp; English.

XX The invention relates to an adeno-associated virus (AAV) clade comprising
 CC at least three AAV members, where each member of the AAV clade is

CC phylogenetically related as determined using a Neighbor-Joining heuristic
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
 CC correction distance measurement of no more than 0.05. The methods and
 CC compositions of the present invention are useful for the prevention
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
 CC colitis. The present sequence represents an adeno-associated virus DNA.

XX
 SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;

Query Match 94.8%; Score 135.6; DB 14; Length 4721;
 Best Local Similarity 97.2%; Pred. No. 2.5e-26;
 Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TTGCCACCTCCCTCTCTGCGGCTCGCTCGGTCGGTGGGGCTGCGGACCAAGGTCCGC 60
 DB 1 TTGGCCACTCCCTCTCTATGCGGCTCGCTCGGTCGGTGGGGCTGCGGACCAAGGTCCGC 60
 QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCAGAGAGGAGTG 120
 DB 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCATAGAGGAGTG 120
 QY 121 GGCACCTCCATCACTAGGGGTA 142
 DB 121 GCCAACTCCATCACTAGGGGTA 142

RESULT 13

AAD00772/c

ID AAD00772 standard; DNA; 4718 BP.

XX AAD00772;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 DNA.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
 KW cap protein; recombinant viral vector; gene delivery; gene therapy;
 KW vaccine; transgene; ss.

XX Adeno associated virus serotype 1.

Key Location/Qualifiers
 repeat_unit 1..143
 /tag= a
 /label= 5' ITR
 /note= "Inverted terminal repeat which is capable of
 forming T-shaped hairpin structure"
 protein_bind 89..110
 /tag= b
 /bound_moiety= "Rep protein"
 misc_feature 124..125
 /tag= c
 /note= "Terminal resolve site (TRS)"
 misc_binding 219..226
 /tag= d
 /bound_moiety= "USF"
 promoter 236..299
 /tag= e
 /label= P5_promoter
 protein_bind 237..245
 /tag= f
 /bound_moiety= "YY1 factor"
 TATA_signal 270..275
 /tag= g
 /label= P5_TATA-Box
 misc_feature 299..306
 /tag= h
 /note= "YY1/p5 RNA"
 CDS 335..2272

```
FT FT /*tag= j
FT FT /product= "Rep 68"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT /note= "The coding region is interrupted by intron"
FT CDS 335..2206
FT FT /*tag= i
FT FT /product= "Rep 78"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT 857..862
FT TATA_signal
FT FT /label= P19_TATA_Box
FT FT 892..883
FT FT /*tag= m
FT FT /note= "P19 RNA"
FT FT 1007..2272
FT CDS
FT FT /*tag= o
FT FT /product= "Rep 40"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT /note= "The coding region is interrupted by intron"
FT CDS 1007..2206
FT FT /*tag= n
FT FT /product= "Rep 52"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT 1836..1841
FT TATA_signal
FT FT /*tag= p
FT FT /label= P40_TATA-BOX
FT FT 1875..1876
FT FT /*tag= g
FT FT /note= "P40 RNA"
FT FT 1924..2220
FT intron
FT FT /*tag= k
FT FT /note= "This region interrupts the coding sequence of Rep
FT FT 68 and Rep 40"
FT CDS 2223..4433
FT FT /*tag= r
FT FT /product= "VP1 protein"
FT FT /note= "Capsid protein"
FT FT 2634..4433
FT CDS
FT FT /*tag= s
FT FT /product= "VP2 protein"
FT FT /note= "Capsid protein"
FT FT /partial
FT FT 2829..4433
FT CDS
FT FT /*tag= t
FT FT /product= "VP3 protein"
FT FT /note= "Capsid protein"
FT FT 4447..4452
FT polyA_signal
FT FT /*tag= u
FT FT 4576..4718
FT repeat_unit
FT FT /*tag= v
FT FT /label= 3' ITR
FT FT /note= "Inverted terminal repeat which is capable of
FT FT forming I-shaped hairpin structure"
FT FT
FT XX WO200028061-A2.
FT PN
FT XX 18-MAY-2000.
FT PD
FT XX 02-NOV-1999; 99WO-US025694.
FT PF
FT XX 05-NOV-1998; 98US-0107114P.
FT PR
FT XX (UYPE-) UNIV PENNSYLVANIA.
FT PA
FT XX Wilson JM, Xiao W;
FT PI
FT XX WPI; 2000-376571/32.
FT DR
FT XX P-PSDB: AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
FT DR AAY71169.
```

```
XX XX Novel adeno-associated virus serotype 1 polynucleotide useful for
PT PT preparation of medicament for delivery of a transgene to a host.
XX XX
XX PS Claim 1; Fig 1; 108pp; English.
XX XX
XX CC The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
CC characterised by two inverted terminal repeats (ITR) and open reading
CC frames for rep and capsid (cap) proteins. The rep reading frame encodes
CC four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
CC frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
CC sequence or its fragments particularly ITRs, rep and cap coding regions,
CC are useful in production of recombinant viral vectors for gene delivery.
CC These vectors can be used as gene therapy vectors, vaccine vectors or
CC antisense delivery vectors. The AAV-1 does not induce the formation of
CC neutralising antibodies specific to any serotype of AAV hence is useful
CC for transforming host cells, and in the preparation of a medicament for
CC the delivery of transgene to a host.
XX XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 83.2%; Score 119; DB 3; Length 4718;
Best Local Similarity 89.5%; Pred. No. 5.8e-22;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGCAGACCAAGGTCGCG 60
DB 4718 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCGGAGAGGAGCTTGC 4659
QY 61 AGACGGCAGAGCTTCTCTTGGCGGCGCCACCGAGCGAGCGAGCGAGAGGAGTG 120
DB 4658 CGTCTCGGAGACCTTTGGTCCGAGGCCGCCAGCGAGCGAGCGAGGAGGAGTG 4599
QY 121 GGCACTCCATCAGGGGTAA 143
DB 4598 GGCACTCCATCAGGGGTAA 4576
RESULT 14
ADE76507/C
ID ADE76507 standard; DNA; 4718 BP.
XX
AC ADE76507;
XX
XX 29-JAN-2004 (first entry)
XX
DE Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
XX
KW adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antiinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease; ds.
XX
OS Adeno-associated virus 1.
XX
XX EP1310571-A2.
XX
XX 14-MAY-2003.
XX
XX 12-NOV-2002; 2002BP-00257826.
XX
XX 13-NOV-2001; 2001US-0350607P.
XX 17-DEC-2001; 2001US-0341117P.
XX 01-MAY-2002; 2002US-0377066P.
XX 05-JUN-2002; 2002US-0386675P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX PA
XX Gao G, Wilson JM, Alvira M;
XX WPI; 2003-450984/43.
XX DR
```

```
XX PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX PS Claim 14; SEQ ID NO 6; 419pp; English.
XX
CC The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antirheumatic, antiarthritic, neuroprotective,
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 83.2%; Score 119; DB 10; Length 4718;
Best Local Similarity 89.5%; Pred. No. 5.8e-22;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTCGCGAGCAAGGTCCGC 60
DB 4718 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTCGCGAGCAAGGTCTGC 4659
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCGAGAGGGAGTG 120
DB 4658 CGTCTGGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGCGCGAGAGGGAGTG 4599
QY 121 GGCAACTCCATCACTAGGGGTAA 143
DB 4598 GGCAACTCCATCACTAGGGGTAA 4576
RESULT 15
ADL13984/c
ID ADL13984 standard; DNA; 4718 BP.
XX AC ADL13984;
XX
XX 06-MAY-2004 (first entry)
XX
XX Adeno-associated virus serotype 1 complete DNA.
XX
XX ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;
XX expression construct; adeno-associated virus;
XX integration efficiency element; inverted terminal repeat; integration;
XX chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;
XX retinoblastoma; inflammatory disease; arthritis;
XX neurodegenerative disease.
XX
XX Adeno-associated virus 1.
XX
XX WO2003087334-A2.
XX
XX 23-OCT-2003.
XX
XX 09-APR-2003; 2003WO-US011191.
XX
XX 09-APR-2002; 2002US-0371044P.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Falck-Pedersen ES, Philpott N;
XX
```

```
DR WPI; 2003-833723/77.
XX
XX New expression construct comprising a nucleic acid sequence encoding an
PT adeno-associated virus integration efficiency element, useful for
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease
PT e.g. arthritis.
XX
XX PS Disclosure; SEQ ID NO 3; 62pp; English.
XX
CC The invention relates to an expression construct comprising a nucleic
CC acid sequence encoding an adeno-associated virus integration efficiency
CC element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV
CC ITRs) and site-specifically integrates into a host cell chromosome when
CC provided to the host cell in conjunction with an AAV Rep protein. The
CC expression construct can be used as a therapeutic factor for treating a
CC mammal for a pathologic state which is cancer, including lung cancer,
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,
CC liver cancer, multiple myeloma, neuroblastoma, skin cancer, stomach cancer,
CC cancer, prostatic cancer, retinoblastoma, ovarian cancer, pancreatic
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic
CC state includes inflammatory disease (arthritis), neurodegenerative
CC disease, a disease of an organ attributed to the presence of increased or
CC decreased level of a particular gene product(s). This sequence
CC corresponds to the AAV serotype 1 complete DNA including the IBE
CC sequence.
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 83.2%; Score 119; DB 10; Length 4718;
Best Local Similarity 89.5%; Pred. No. 5.8e-22;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTCGCGAGCAAGGTCCGC 60
DB 4718 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTCGCGAGCAAGGTCTGC 4659
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCGAGAGGGAGTG 120
DB 4658 CGTCTGGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGCGCGAGAGGGAGTG 4599
QY 121 GGCAACTCCATCACTAGGGGTAA 143
DB 4598 GGCAACTCCATCACTAGGGGTAA 4576
```

Search completed: November 28, 2005, 19:55:37
Job time : 69.9124 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 454.049 Seconds
(without alignments)
14735.325 Million cell updates/sec

Title: US-10-696-900-1_COPY_1_143

Perfect score: 143

Sequence: 1 ttgccacctcctctctgcg.....aactccatcactaggggtaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	38.2	26.7	763	2	BG490707
c 2	37.8	26.4	1021	5	B0838990
c 3	37.4	26.2	644	2	BB632576
c 4	36.2	25.3	920	2	BG764368
c 5	36.2	25.3	932	10	CNS0072Q
c 6	36.2	25.3	1021	5	B0838990
c 7	35.8	25.0	604	3	B1531460
c 8	35.8	25.0	1074	10	CL507458
c 9	35.6	24.9	578	2	BB651469
c 10	35.4	24.8	860	2	BF578052
c 11	35.4	24.8	1123	10	AG080476
c 12	35.2	24.6	2024	10	AG074763
c 13	35	24.5	724	10	CL751558
c 14	35	24.5	844	10	CNS0052P
c 15	35	24.5	1303	9	BZ570288
c 16	34.8	24.3	676	9	BZ414423
c 17	34.8	24.3	926	8	CV775269
c 18	34.6	24.2	467	9	CC557657
c 19	34.6	24.2	617	7	CN038418
c 20	34.6	24.2	637	7	CN045740
c 21	34.6	24.2	2088	10	CL507788
c 22	34.4	24.1	318	7	CV352119

c 23	34.4	24.1	745	2	BG837739	BG837739 Zml0_04hl
c 24	34.4	24.1	890	10	AG127787	AG127787 Pan trogl
c 25	34.2	23.9	364	5	BY014749	BY014749 BY014749
c 26	34.2	23.9	468	8	CX358645	CX358645 JGI_XZT15
c 27	34.2	23.9	1367	10	CW943432	CW943432 TcB25.2.A
c 28	34	23.8	343	5	BY181775	BY181775 BY181775
c 29	34	23.8	450	5	BY250974	BY250974 BY250974
c 30	34	23.8	623	2	BB613613	BB613613 BB613613
c 31	34	23.8	631	5	BY723537	BY723537 BY723537
c 32	34	23.8	639	2	BB651977	BB651977 BB651977
c 33	34	23.8	643	2	BB652152	BB652152 BB652152
c 34	34	23.8	657	2	BB651054	BB651054 BB651054
c 35	34	23.8	661	2	BB621081	BB621081 BB621081
c 36	34	23.8	678	5	BY726305	BY726305 BY726305
c 37	34	23.8	795	9	AQ331299	AQ331299 nxb0049M
c 38	34	23.8	838	8	CV780024	CV780024 FGAS07443
c 39	34	23.8	1364	5	BU543298	BU543298 AGENCOURT
c 40	34	23.8	1767	4	AK082810	AK082810 Mus muscu
c 41	34	23.8	2268	4	AK082478	AK082478 Mus muscu
c 42	34	23.8	2708	4	AK038836	AK038836 Mus muscu
c 43	34	23.8	3798	4	AK031553	AK031553 Mus muscu
c 44	34	23.8	3994	4	AK029139	AK029139 Mus muscu
c 45	33.8	23.6	395	10	CW754890	CW754890 OG_BB005

ALIGNMENTS

RESULT 1
BG490707/c
LOCUS
DEFINITION
602520211F1 NIH_MGC_18 Homo sapiens CDNA clone IMAGE:4638699 5',
mRNA sequence.
ACCESSION
BG490707
VERSION
BG490707.1
KEYWORDS
GI:13452217
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 763)
NIH-MGC http://imgc.ncbi.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTF/Gazdar
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1401 row: e column: 04
High quality sequence stop: 182.

FEATURES
source
1..763
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4638699"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/notes="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

/clone_lib="RIKEN full-length enriched, adult male
hypothenamus"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATTCTCGAGTTAATTAATTAATCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

ORIGIN

Query Match 26.2%; Score 37.4; DB 2; Length 644;
Best Local Similarity 64.4%; Pred. No. 17;
Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 11 CCTCTCTGGCGCTCGTCTGGTGGGCGCTGCGGACCAAGTCCGACAGCGCAGA 70
Db 535 CCTCTCGCGCGCTGCGCAGGTCTGTGGCGCGGCGGAGACAGCAGCGCCATG 594
QY 71 GCTCTCTGCGCGCGCCACCGACGCG 97
Db 595 CCTCTACTTTGACGTGCCCCCGCGTG 621

RESULT 4

BG764368
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI720 row: d column: 16
High quality sequence stop: 748.

FEATURES

source
1. .920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4861407"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:
GCCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Super-script
II RT (Life Technologies). Note: this is a NIH_MGC
Library. |

ORIGIN

Query Match 25.3%; Score 36.2; DB 2; Length 920;
Best Local Similarity 56.2%; Pred. No. 34;
Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 20 GCGCTCGTCTCGTCTGGTGGGCGCTGCGGACCAAGTCCGACAGCGGAGCTCTGCTC 79
Db 602 GCGCTGGGCGAGGCTTAGGGTCTGTGGGCGCAACCCCTGGGTGGCAAGCCGCTCG 661
QY 80 TCCCGCGCCCGACGAGCGAGCGCGCGACAGAGGAGTGGGCACTCATCTACCTAGGG 139
Db 662 TCACGGCTCTACCTGTGTAGGCGCCCTGTGACACAGGAGACACCTCTTGTCTCATGG 721
QY 140 G 140
Db 722 G 722

RESULT 5

CNS0072Q
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
932 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BAC14B09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL066742
GI:4945205
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 932)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Hammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1. .932
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BAC14B09"
/clone_lib="RPCI-98"
/note="end : T7"
ORIGIN
Query Match 25.3%; Score 36.2; DB 10; Length 932;
Best Local Similarity 36.3%; Pred. No. 34;

```

Matches 41; Conservative 32; Mismatches 40; Indels 0; Gaps 0;

QY 4 CCACATCCCTCTCTGCGGCTCGCTCGCTCGCTCGGACCAAGGTCCGCAGA 63
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 725 SBCGSCCCSCSCCCSCCCSCGSGGCGGCGGCGGCGGCGGCGGCGG 784
QY 64 CGGAGAGCTCTGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 116
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 785 CGGSCGSGGCGGSGGCGGSGGSGGCGGSGGCGGSGGSGGCGGSGG 837

RESULT 6
LOCUS BU838990
DEFINITION AGENCOURT_8209898 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6258289
5', mRNA sequence.
ACCESSION BU838990
VERSION BU838990.1 GI:24023385
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bs-r@mail.nih.gov
Tissue Procurement: DCTB/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2413 row: p column: 02
High quality sequence start: 36
High quality sequence stop: 231.

FEATURES
Location/Qualifiers
1..1021
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6258289"
/tissue_type="melanotic melanoma, cell line"
/lab_host="NIH MGC_112"
/clone_lib="NIH MGC_112"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 25.3%; Score 36.2; DB 5; Length 1021;
Best Local Similarity 57.5%; Pred. No. 34;
Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 4 CCACATCCCTCTCTGCGGCTCGCTCGCTCGCTCGGACCAAGGTCCGCAGA 63
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 791 CCCCCACCCACCCCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 850
QY 64 CGGAGAGCTCTGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 116
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 851 CCCCCGACCGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 903

RESULT 7

```

```

BI531460/C
LOCUS BI531460
DEFINITION 102413H09.y1 C. reinhardtii CC-1690, normalized, Lambda Zap II
Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BI531460
VERSION BI531460.1 GI:15372034
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
REFERENCE Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C.,
AUTHORS Lefebvre,P., McDermott,J.P., Shreager,J., Silflow,C. and Stern,D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1024b
JOURNAL Unpublished (2001)
COMMENT Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
Location/Qualifiers
1..604
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 2lgr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
McDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light, TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with Exsist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

ORIGIN
Query Match 25.0%; Score 35.8; DB 3; Length 604;
Best Local Similarity 59.2%; Pred. No. 43;
Matches 61; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 8 CTCCTCTCTGCGGCTCGCTCGCTCGGCTCGGACCAAGGTCCGCAGCGC 67
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 382 CTTCTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 323
QY 68 AGAGCTCTGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 110
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 322 AGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 280

RESULT 8
LOCUS CL507458
DEFINITION SAIL_781_E02.v2 SAIL Collection Arabidopsis thaliana genomic clone
SAIL_781_E02.v2, genomic survey sequence.
ACCESSION CL507458
VERSION CL507458.1 GI:46004778
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Arabidopsis thaliana
AUTHORS Arabidopsis thaliana
TITLE Arabidopsis thaliana
JOURNAL Arabidopsis thaliana
COMMENT Arabidopsis thaliana

```

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS Sessions, A., Burke, E., Preating, G., Aux G., McEliver, J., Patton, D., Dietrich, B., Ho, P., Bawaden, J., Ko, C., Clarke, J.D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmery, B., Mittel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.
TITLE A high-throughput Arabidopsis reverse genetics system
JOURNAL Plant Cell 14 (12), 2985-2994 (2002)
PUBMED 12468722

COMMENT Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS934967: T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not single contiguous sequences.
Class: TDNA tagged.

FEATURES
source Location/Qualifiers
1..1074
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL 781 E02.v2"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN
Query Match 25.0%; Score 35.8; DB 10; Length 1074;
Best Local Similarity 57.7%; Pred. No. 43;
Matches 64; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 13 TCTCTCGCGCTCGCTCGCTCGCTGGGGCTGGGACCAAGTCCGCGAGCGGCGAGCG 72
DB 865 TCGCGCGCGCGCGCGAGATCGGGCGCGAGCGGCGGAGCAGTTCGGGCGCGCAGCA 924
QY 73 TTTGTTCTCGCGCCCGCCCGAGCGAGCGAGCGCGCGAGAGAGGAGTGGGC 123
DB 925 GCTCGCGCGCGCGCGCGCGCGGTAGAGCGCGCGCGTACGACGACGCGAC 975

RESULT 9
BB651469 578 bp mRNA linear EST 26-OCT-2001
LOCUS BB651469 RIKEN full-length enriched, 0 day neonate cerebellum Mus
DEFINITION musculus cDNA clone C230083H05 5', mRNA sequence.
ACCESSION BB651469
VERSION BB651469.1 GI:16485722
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 578)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ono, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)

CONTACT: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES
source Location/Qualifiers
1..578
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="C230083H05"
/tissue_type="cerebellum"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate cerebellum"
/note="Site 1: Sall; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 479.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAATAAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN
Query Match 24.9%; Score 35.6; DB 2; Length 578;
Best Local Similarity 64.6%; Pred. No. 49;
Matches 53; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 11 CCTCTCTCGCGCTCGCTCGCTCGGTGGGCGCTCGGACCAAGGTCCGCGAGCGGCGAGA 70
DB 22 CCTCTCTCGCGCGCTCGCGCAGGTTCGGGCGCGCGGCGAGAGACACGACGCGCATG 81
QY 71 GCTCTGCTCTCGCGCGCCGCC 92
DB 82 CTCTACTTTGCGTGCACCC 103

RESULT 10
BF578052

AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

COMMENT

Tel: 81-45-503-9111, Fax: 81-45-503-9112
Clones are derived from the chimpanzee

-9170)
anzee BAC library PTB This BAC ends

Query Match	24.5%;	Score 35;	DB 10;	Length 724;
Best Local Similarity	55.3%;	Pred. No. 70;		
Matches	68;	Conservative	0;	Mismatches 55; Indels 0; Gaps 0;
QY	1	TTGCCCACTCCCTCTCTGTCGGCGCTCGCTCGCTCGGTGGGGCTTCGCGGACCAAGGTCGCG	60	
Ddb	438	TTATCTCTCTCTCTACTCTCTCTCTCTCACCAGAGCGCTGCGAGGAGCGCGGCTCGG	497	
QY	61	AGACGGCAGAGCTCTGTCTCTGTCGCGGCGCCACCAGAGCGAGCGGCGCAGAGGAGTG	120	
Ddb	498	CGACGAAGCGGCGCTCCGCTGGCGCGGTACTACGGGAGCGCGCGGTCAATTACGAGGAG	557	
QY	121	GGC 123		
Ddb	558	GGC 560		
RESULT 14				
CNS0052P				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
ORIGIN				
Query Match	24.5%;	Score 35;	DB 10;	Length 844;
Best Local Similarity	21.5%;	Pred. No. 70;		
Matches	26;	Conservative	61;	Mismatches 34; Indels 0; Gaps 0;
QY	3	GCCACTCCCTCTCTCGCGGCTCGCTCGCTCGGTGGGCTTCGGACCAAGGTCGCGAG	62	
Ddb	490	SMCCSMCCSVSSVCAYCCSSGVSRSVAVCGSVGRVGGSSRGRAGRSSRGGGSSVSGW	549	
QY	63	ACGGCAGAGCTCTGTCTGCCGCGCCACCAGCGGAGCGCGCAGAGGAGGAGTGCGG	122	

Db 550 8888SVGMGCACSSASSVSCBSSVSVSSGBVSRGCRVCVGGVGGSRVSSCSGSS 609
 QY 123 C 123
 Db 610 S 610

RESULT 15

BZ570288
 LOCUS BZ570288 1303 bp DNA linear GSS 17-DEC-2002
 DEFINITION mesh2_1301.x1 mesh Pseudomonas aeruginosa genomic clone mesh2_1301,
 genomic survey sequence.

ACCESSION BZ570288

VERSION BZ570288.1 GI:27205349

KEYWORDS GSS.

SOURCE Pseudomonas aeruginosa

ORGANISM Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 1303)

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

JOURNAL Contact: Chris K. Raymond

COMMENT Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

FEATURES

source Location/Qualifiers
 1..1303
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="MSH"
 /db_xref="taxon:287"
 /clone="mesh2_1301"
 /clone_lib="mesh"
 /note="Environmental isolate. Whole genomic shotgun
 library."

ORIGIN

Query Match 24.5%; Score 35; DB 9; Length 1303;
 Best Local Similarity 55.3%; Pred. NO. 69;
 Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 13 TCTCTGGCGCTCGCTCGCTCGCTGGGGCTCGGACCAAGGTCCGACAGCGGAGGC 72
 Db 550 TATCGGACGGCGCCACGCTCGAACAGCCCCGGGGACCCCGGGCTCGGACGCAAAAGG 609
 QY 73 TGTGTTCTCCGCGCCCGCCAGCGAGCGGAGCGGAGAGGAGGTGGGCAACTCCATC 132
 Db 610 CCGCCCGAGCGGGCTTGGGACAGCGGACGACACCCGAGGGGTAAAGATGGGCCCATG 569
 QY 133 ACT 135
 Db 670 ACT 672

Search completed: November 29, 2005, 14:17:49
 Job time : 458.049 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 08:46:14 ; Search time 66.9124 Seconds
(without alignments)
14243.261 Million cell updates/sec

Title: US-10-696-900-1_COPY_4576_4718

Perfect score: 143

Sequence: 1 ttacccttagtgatggagtt.....cgcagaggggagtgggcaaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*
- 14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	4718	3	AAD00772
2	143	100.0	4718	10	ADG76507
3	143	100.0	4718	10	ADL13984
4	143	100.0	4718	12	ADG39758
5	135.6	94.8	4721	10	ADG397502
6	135.6	94.8	4721	12	ADG39764
7	135.6	94.8	4721	14	ADG39764
8	135.6	94.8	4721	14	ADG39764
9	125	87.4	4718	3	AAD00781
10	121.8	85.2	4683	4	AAF23749
11	121.8	85.2	4683	10	ADL13983
12	121.8	85.2	4683	12	ADG39763
13	119	83.2	4718	3	AAD00781
14	119	83.2	4718	3	AAD00772
15	119	83.2	4718	10	ADG76507
16	119	83.2	4718	10	ADL13984
17	119	83.2	4718	12	ADG39758
18	114.8	80.3	4683	4	AAF23749
19	114.8	80.3	4683	10	ADL13983

20	114.8	80.3	4683	12	ADG39763	Adg39763 AAV-6 gen
21	113.2	79.2	4726	10	ADG76509	Adg76509 Adeno-ass
22	113.2	79.2	4726	10	ADL13985	Adl13985 Adeno-ass
23	113.2	79.2	4726	12	ADG39759	Adg39759 AAV-3A ge
24	111.6	78.0	4721	10	ADG76502	Adg76502 Adeno-ass
25	111.6	78.0	4721	12	ADG39764	Adg39764 AAV-7 gen
26	111.6	78.0	4721	14	ADG76509	Adg76509 Nucleotid
27	111.6	78.0	4721	14	ADG76509	Adg76509 Nucleotid
28	107.4	75.1	144	9	ACF35877	Adf35877 AAV-2 rig
29	107.4	75.1	145	2	AA02990	AA02990 Adeno-ass
30	107.4	75.1	145	6	ABA02990	ABA02990 Adeno-ass
31	107.4	75.1	145	13	ADM93254	Adm93254 Inverted
32	107.4	75.1	165	2	AAQ66769	AAQ66769 Double-D
33	107.4	75.1	165	2	AA049462	AA049462 Adeno ass
34	107.4	75.1	165	4	AA003535	AA003535 Inverted
35	107.4	75.1	165	8	ABV77279	ABV77279 Nucleotid
36	107.4	75.1	192	2	AA033384	AA033384 Right-han
37	107.4	75.1	207	6	AA044621	AA044621 165 bp en
38	107.4	75.1	272	3	AA047166	AA047166 Head-to-t
39	107.4	75.1	272	3	AA003534	AA003534 AAV circ
40	107.4	75.1	4675	4	AAH41481	AAH41481 Adeno-ass
41	107.4	75.1	4675	4	AAF89931	AAF89931 Nucleotid
42	107.4	75.1	4675	6	ABA02989	ABA02989 Adeno-ass
43	107.4	75.1	4675	6	ABS69879	ABS69879 Human ade
44	107.4	75.1	4675	6	ABS69880	ABS69880 Human ade
45	107.4	75.1	4675	10	ADG76508	Adg76508 Adeno-ass

ALIGNMENTS

RESULT 1

AAD00772

ID AAD00772 standard; DNA; 4718 BP.

XX AAD00772;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 DNA.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

```

FT FT /*tag= h
FT FT /note= "VP1/p5 RNA"
FT FT 335..2272
FT FT /tag= j
FT FT /product= "Rep 68"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT /note= "The coding region is interrupted by intron"
FT FT 335..2206
FT FT /tag= i
FT FT /product= "Rep 78"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT 857..862
FT FT /label= P19_TATA_Box
FT FT 882..883
FT FT /tag= m
FT FT /note= "P19 RNA"
FT FT 1007..2272
FT FT /tag= o
FT FT /product= "Rep 40"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT /note= "The coding region is interrupted by intron"
FT FT 1007..2206
FT FT /tag= n
FT FT /product= "Rep 52"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT 1836..1841
FT FT /tag= p
FT FT /label= P40_TATA_BOX
FT FT 1875..1876
FT FT /tag= q
FT FT /note= "P40 RNA"
FT FT 1924..2220
FT FT /tag= k
FT FT /note= "This region interrupts the coding sequence of Rep
FT FT 68 and Rep 40"
FT FT 2223..4433
FT FT /tag= r
FT FT /product= "VP1 protein"
FT FT /note= "Capsid protein"
FT FT 2634..4433
FT FT /tag= s
FT FT /product= "VP2 protein"
FT FT /note= "Capsid protein"
FT FT /partial
FT FT 2829..4433
FT FT /tag= t
FT FT /product= "VP3 protein"
FT FT /note= "Capsid protein"
FT FT 4447..4452
FT FT /tag= u
FT FT /tag= v
FT FT 4576..4718
FT FT /tag= v
FT FT /label= 3' _ITR
FT FT /note= "Inverted terminal repeat which is capable of
FT FT forming T-shaped hairpin structure"
FT FT
XX XX WO200028061-A2.
XX XX
XX XX 18-MAY-2000.
XX XX
XX XX 02-NOV-1999; 99WO-US025694.
XX XX
XX XX 05-NOV-1998; 98US-0107114P.
XX XX
XX XX (UYPE-) UNIV PENNSYLVANIA.
XX XX
XX XX Wilson JM, Xiao W;
XX XX

```

```

DR WPI; 2000-376571/32.
DR P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
DR AAY71169.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host.
XX
PS Claim 1; Fig 1; 108pp; English.
XX
CC The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
CC characterised by two inverted terminal repeats (ITR) and open reading
CC frames for rep and capsid (cap) proteins. The rep reading frame encodes
CC four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
CC frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
CC sequence or its fragments particularly ITRs, rep and cap coding regions,
CC are useful in production of recombinant viral vectors for gene delivery.
CC These vectors can be used as gene therapy vectors, vaccine vectors or
CC antisense delivery vectors. The AAV-1 does not induce the formation of
CC neutralising antibodies specific to any serotype of AAV hence is useful
CC for transforming host cells, and in the preparation of a medicament for
CC the delivery of transgene to a host
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 3; Length 4718;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGTGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGCTCGTGGGCC 60
Db 4576 TTACCCCTAGTGTGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGCTCGTGGGCC 4635
QY 61 TCGCGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGA 120
Db 4636 TCGCGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
Db 4696 GCGCGCAGAGAGGAGTGGGCAA 4718
RESULT 2
ADE76507
ID ADE76507 standard; DNA; 4718 BP.
XX
AC ADE76507;
XX
DT 29-JAN-2004 (first entry)
XX
DE Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
XX
KW adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antiinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
OS Adeno-associated virus 1.
XX
PN EP1310571-A2.
XX
PD 14-MAY-2003.
XX
XX 12-NOV-2002; 2002EP-00257826.
XX
XX 13-NOV-2001; 2001US-0350607P.
XX 17-DEC-2001; 2001US-0341117P.
XX 01-MAY-2002; 2002US-0377066P.
XX 05-JUN-2002; 2002US-0386675P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX

```


PI Gao G, Wilson JM, Alvira M;
XX WPI; 2003-450984/43.
XX
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX
XX Claim 14; SEQ ID NO 6; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antirheumatic, antialarthritis, neuroprotective,
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences such as e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 143; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCTCTCTCGCGCTCGCTCGCTCGGTGGGGCC 60
DB 4576 TTACCCCTAGTGATGAGTTGCCACTCTCTCTCGCGCTCGCTCGGTGGGGCC 4635
QY 61 TCGCGACCAAGGTCGCGAGCGGAGAGCTCTCTCTCGCGGCCCCACCGAGCGAGCGA 120
DB 4636 TCGCGACCAAGGTCGCGAGCGGAGAGCTCTCTCTCGCGGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGGTGGGCAA 143
DB 4696 GCGCGCAGAGGAGGTGGGCAA 4718
RESULT 3
ADL13984
ID ADL13984 standard; DNA; 4718 BP.
XX
XX ADL13984;
XX
XX 06-MAY-2004 (first entry)
XX
XX Adeno-associated virus serotype 1 complete DNA.
DE
XX
XX aa; cytostatic; neuroprotective; antiinflammatory; gene therapy;
KW expression construct; adeno-associated virus;
KW integration efficiency element; inverted terminal repeat; integration;
KW chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;
KW retinoblastoma; inflammatory disease; arthritis;
KW neurodegenerative disease.
XX
XX Adeno-associated virus 1.
OS
XX WO2003087334-A2.
XX
XX 23-OCT-2003.
XX
XX 09-APR-2003; 2003WO-US011191.
XX
XX 09-APR-2002; 2002US-0371044P.
XX
XX (CORR) CORNELL RES FOUND INC.

XX Falck-Pedersen ES, Philpott N;
PI
XX WPI; 2003-833723/77.
XX
XX New expression construct comprising a nucleic acid sequence encoding an
PT adeno-associated virus integration efficiency element, useful for
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease
PT e.g. arthritis.
XX
XX Disclosure; SEQ ID NO 3; 62pp; English.
XX
XX The invention relates to an expression construct comprising a nucleic
CC acid sequence encoding an adeno-associated virus integration efficiency
CC element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV
CC ITRs) and site-specifically integrates into a host cell chromosome when
CC provided to the host cell in conjunction with an AAV Rep protein. The
CC expression construct can be used as a therapeutic factor for treating a
CC mammal for a pathologic state which is cancer, including lung cancer,
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,
CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic
CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic
CC state includes inflammatory disease (arthritis), neurodegenerative
CC disease, a disease of an organ attributed to the presence of increased or
CC decreased level of a particular gene product(s). This sequence
CC corresponds to the AAV serotype 1 complete DNA including the IEE
CC sequence.
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 143; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCTCTCTCGCGCTCGCTCGGTGGGGCC 60
DB 4576 TTACCCCTAGTGATGAGTTGCCACTCTCTCTCGCGCTCGCTCGGTGGGGCC 4635
QY 61 TCGCGACCAAGGTCGCGAGCGGAGAGCTCTCTCTCGCGGCCCCACCGAGCGAGCGA 120
DB 4636 TCGCGACCAAGGTCGCGAGCGGAGAGCTCTCTCTCGCGGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGGTGGGCAA 143
DB 4696 GCGCGCAGAGGAGGTGGGCAA 4718
RESULT 4
ADG39758
ID ADG39758 standard; DNA; 4718 BP.
XX
XX ADG39758;
XX
XX 11-MAR-2004 (first entry)
XX
XX AAV-1 genomic DNA sequence SEQ ID NO:20.
DE
XX
XX parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
KW recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
XX Adeno-associated virus 1.
OS
XX WO2003104392-A2.
XX
XX 18-DEC-2003.
XX
XX 02-DEC-2002; 2002WO-US038423.
XX
XX

XX 18-DEC-2001; 2001US-0341919P.
XX (UYN-) UNIV NORTH CAROLINA.
XX Samulski RJ, Rabinowitz JE;
XX MPI; 2004-062324/06.
XX
XX New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
XX Disclosure; SEQ ID NO 20; 115pp; English.
XX
XX The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (I); (2) a cell comprising (I), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprises sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 143; DB 12; Length 4718;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCC 60
DB 4576 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCC 4635
QY 61 TCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTTCCCGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTTCCCGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGGGAGTGGGCAA 4718
RESULT 5
ADE76502
ID ADE76502 standard; DNA; 4721 BP.
XX
XX ADE76502;
AC
XX 29-JAN-2004 (first entry)
DT
XX Adeno-associated virus (AAV) related DNA, SEQ ID No 1.
DE
XX adeno-associated virus; AAV; cycostatic; antipsoriatic; antirheumatic;
KW antithyroid; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antiinflammatory; gene therapy; vaccine;
KW

KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
XX Adeno-associated virus 7.
OS
XX EPI310571-A2.
PN
XX 14-MAY-2003.
PD
XX 12-NOV-2002; 2002EP-00257826.
PF
XX 13-NOV-2001; 2001US-0350607P.
PR
XX 17-DEC-2001; 2001US-0341117P.
PR
XX 01-MAY-2002; 2002US-0377066P.
PR
XX 05-JUN-2002; 2002US-0386675P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
PA
XX Gao G, Wilson JM, Alvira M;
XX WPI; 2003-450984/43.
XX P-PSDB; ADE76503.
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX Claim 17; SEQ ID NO 1; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antithyroid, antidiabetic, antirheumatic, antitumor,
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
XX Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;
SQ
Query Match 94.8%; Score 135.6; DB 10; Length 4721;
Best Local Similarity 97.2%; Pred. No. 3.1e-26;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 TACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCC 61
DB 4580 TACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCC 4639
QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTTCCCGCCCCACCGAGCGAGCGAG 121
DB 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTTCCCGCCCCACCGAGCGAGCGAG 4699
QY 122 GCGGACAGAGGGAGTGGGCAA 143
DB 4700 GCGGACAGAGGGAGTGGGCAA 4721
RESULT 6
ADG39764
ID ADG39764 standard; DNA; 4721 BP.
XX
XX ADG39764;
AC
XX 11-MAR-2004 (first entry)
DT
XX

DE AAV-7 genomic DNA sequence SEQ ID NO:26.
XX parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
KW recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
XX
OS Adeno-associated virus 7.
XX
XX WO2003104392-A2.
XX
XX 18-DEC-2003.
XX
XX 02-DEC-2002; 2002WO-US038423.
XX
XX 18-DEC-2001; 2001US-0341919P.
XX
XX (UYNC-) UNIV NORTH CAROLINA.
XX
XX Samulski RJ, Rabinowitz JE;
XX
XX WPI; 2004-062324/06.
XX
XX New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
XX
XX Disclosure; SEQ ID NO 26; 115pp; English.
XX
XX The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (I); (2) a cell comprising (I), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
XX Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;
SQ
Query Match 94.8%; Score 135.6; DB 12; Length 4721;
Best Local Similarity 97.2%; Pred. No. 3.1e-26;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 TACCCCTAGTGATGGAGTTGCCACATCCCTCTCTGCGCGCTCGCTCGGTGGGCGCT 61
DB 4580 TACCCCTAGTGATGGAGTTGGCCACATCCCTCTATGCGCGCTCGCTCGGTGGGCGCT 4639
QY 62 GCGGACCAAGGTCGCGACAGCGCAGAGTCTCTCTGCGCGGCCACCGAGCGAGCGAG 121
DB 4640 GCGGACCAAGGTCGCGACAGCGCAGAGTCTCTCTGCGCGGCCACCGAGCGAGCGAG 4699
QY 122 CGCGCAGAGAGGAGTGGGCAA 143
DB 4700 CGCGCATAGAGGAGTGGGCAA 4721

RESULT 7
ADV67509
ID ADV67509 standard; DNA; 4721 BP.
XX
XX ADV67509;
AC
XX
XX 10-MAR-2005 (first entry)
DT
XX
XX Nucleotide sequence of AAV serotype 7.
DE
XX
XX antiarteriosclerotic; antilipemic; gene therapy; cholesterol;
KW apolipoprotein E; apoE; apolipoprotein A; apoA; atherosclerosis;
KW lipoprotein defect; ss.
XX
XX Adeno-associated virus.
OS
XX
XX Key Location/Qualifiers
FH repeat_region 1..107
FT /*tag= a
FT /note= "5', ITR"
FT misc_feature 825..3049
FT /*tag= b
FT /note= "encodes vp1"
FT misc_feature 1234..3049
FT /*tag= c
FT /note= "encodes vp2"
FT misc_feature 1434..3049
FT /*tag= d
FT /note= "encodes vp3"
FT repeat_region 4704..4721
FT /*tag= e
FT /note= "3', ITR"
XX
XX WO2004108922-A2.
XX
XX 16-DEC-2004.
XX
XX 23-APR-2004; 2004WO-US010965.
XX
XX 25-APR-2003; 2003US-0465293P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Rader DJ, Wilson JM;
XX
XX WPI; 2005-031700/03.
XX
XX Lowering total cholesterol levels and treating atherosclerosis in a
PT subject comprises delivering a recombinant adeno-associated virus (AAV)
PT comprising an AAV serotype capsid protein or a gene encoding human
PT apolipoprotein E (apoE) or apoA.
XX
XX Disclosure; SEQ ID NO 7; 69pp; English.
XX
XX The specification describes a method for lowering total cholesterol
XX levels in a subject. The method comprises delivering to the subject a
XX recombinant adeno-associated virus (AAV) comprising a gene encoding a
XX human apolipoprotein E (apoE) or apoA under the control of a regulatory
XX control sequence which directs expression of the gene. The recombinant
XX AAV also comprises a capsid protein selected from an AAV serotype (e.g.
XX AAV7 or AAV8) which preferentially expresses high levels of transgene in
XX live. A therapeutically effective amount of apoE or apoA expression is
XX obtained upon delivery of low dose of AAV. The method of the invention is
XX useful for lowering total cholesterol levels in a subject, e.g. for
XX treating atherosclerosis, and for correcting defects in lipoprotein. The
XX present sequence represents AAV serotype 7 (AAV7), which may be used to
XX produce recombinant AAV vectors of the invention.
XX
XX Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;
SQ
Query Match 94.8%; Score 135.6; DB 14; Length 4721;
Best Local Similarity 97.2%; Pred. No. 3.1e-26;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Matches	138;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	2	TACCCCTAGTGTGGAGTTGCCACTCCCTCTCTTCGCGCGCTCGCTCGCTCGGTGGGCGCT	61						
Db	4580	TACCCCTAGTGTGGAGTTGCCACTCCCTCTATGCGGCTCGCTCGCTCGGTGGGCGCT	4639						
Qy	62	CGCGACCAAAAGGTCGCGAGCGGAGAGCTCTGTCTTCGCGGCCCAACCGAGCGAGCGAG	121						
Db	4640	CGCGACCAAAAGGTCGCGAGCGGAGAGCTCTGTCTTCGCGGCCCAACCGAGCGAGCGAG	4699						
Qy	122	CGCGCAGAGAGGGAGTGGGCAA	143						
Db	4700	CGCGCATAGAGGGAGTGGGCAA	4721						

RESULT 8
ADZ27030
ID ADZ27030 standard; DNA; 4721 BP.
XX AC
XX AC ADZ27030;
XX AC
XX DT 30-JUN-2005 (first entry)
XX DE
XX DE Adueno-associated virus DNA SEQ ID NO 180.
XX KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antitumor; dermatological; ds.
XX OS
XX OS Adueno-associated virus.
XX PN W02005033321-A2.
XX PD
XX PD 14-APR-2005.
XX PF
XX PF 30-SEP-2004; 2004WO-US028817.
XX PR
XX PR 30-SEP-2003; 2003US-0508226P.
XX PR 29-APR-2004; 2004US-0566546P.
XX PA (UYPE-) UNIV PENNSYLVANIA.
XX PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX MPI; 2005-285437/29.
XX DR
XX PT New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX PS
XX PS Claim 12; SEQ ID NO 180; 569pp; English.
XX CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX SQ
XX SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;
Query Match 94.8%; Score 135.6; DB 14; Length 4721;
Best Local Similarity 97.2%; Pred. No. 3.1e-26;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CC viral vector. The ITR forms a T-shaped hairpin structure
XX
SQ Sequence 143 BP; 24 A; 50 C; 48 G; 21 T; 0 U; 0 Other;

Query Match 87.4%; Score 125; DB 3; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 78
DB 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 60

QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 138
DB 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120

QY 139 GGCNA 143
DB 121 GGCNA 125

RESULT 10
AAF23749/c
ID AAF23749 standard; DNA; 4683 BP.

AC AAF23749;

DT 28-MAR-2001 (first entry)

XX AAV6 DNA sequence.

XX AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassaemia;
KW blood clotting disorder; diabetes; ss.

XX Adeno associated virus.

XX US6156303-A.

XX 05-DEC-2000.

XX 11-JUN-1997; 97US-00873168.

XX 11-JUN-1997; 97US-00873168.

XX (UNIW) UNIV WASHINGTON.

XX Russell DW, Rutledge EA;

XX WPI; 2001-060164/07.

XX Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassaemia and diabetes.

XX Claim 1; Fig 1; 50pp; English.

XX The present invention relates to adeno-associated virus serotypes. The
CC present sequence is the DNA sequence of one such serotype (AAV6). AAV6
CC can be used to construct AAV viral vectors for use in gene therapy for a
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell

XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;

Query Match 85.2%; Score 121.8; DB 4; Length 4683;
Best Local Similarity 98.4%; Pred. No. 1.3e-22;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 78
DB 4683 TTGCCCACTCCCTCTATCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAGGTCCGC 4624

QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 138
DB 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4564

QY 139 GGCNA 143
DB 4563 GGCNA 4559

ADL13983/c

ID ADL13983 standard; DNA; 4683 BP.

XX ADL13983;

XX 06-MAY-2004 (first entry)

XX Adeno-associated virus serotype 6 complete DNA.

XX ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;
KW expression construct; adeno-associated virus;
KW integration efficiency element; inverted terminal repeat; integration;
KW chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;
KW retinoblastoma; inflammatory disease; arthritis;
KW neurodegenerative disease.

XX Adeno-associated virus 6.

XX WO2003087334-A2.

XX 23-OCT-2003.

XX 09-APR-2003; 2003WO-US011191.

XX 09-APR-2002; 2002US-0371044P.

XX (CORR) CORNELL RES FOUND INC.

XX Falck-Pedersen ES, Philpott N;

XX WPI; 2003-833723/77.

XX New expression construct comprising a nucleic acid sequence encoding an
PT adeno-associated virus integration efficiency element, useful for
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease
PT e.g. arthritis.

XX Disclosure; SEQ ID NO 2; 62pp; English.

XX The invention relates to an expression construct comprising a nucleic
CC acid sequence encoding an adeno-associated virus integration efficiency
CC element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV
CC ITRs) and site-specifically integrates into a host cell chromosome when
CC provided to the host cell in conjunction with an AAV Rep protein. The
CC expression construct can be used as a therapeutic factor for treating a
CC mammal for a pathologic state which is cancer, including lung cancer,
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,
CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic
CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic
CC state includes inflammatory disease (arthritis), neurodegenerative
CC disease, a disease of an organ attributed to the presence of increased or
CC decreased level of a particular gene product(s). This sequence
CC corresponds to the AAV serotype 6 complete DNA including the IEE
CC sequence.

XX

SQ Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
 Query Match 85.2%; Score 121.8; DB 10; Length 4683;
 Best Local Similarity 98.4%; Pred. No. 1.3e-22;
 Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 19 TTGCCCACTCCTCTCTGCGGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 78
 DB 4683 TTGCCCACTCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 4624
 QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGCGCAGAGAGGAGTG 138
 DB 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGCGCAGAGAGGAGTG 4564
 QY 139 GGCAA 143
 DB 4563 GGCAA 4559
 RESULT 12
 ADG39763/C
 ID ADG39763 standard; DNA; 4683 BP.
 AC ADG39763;
 DT 11-MAR-2004 (first entry)
 XX AAV-6 genomic DNA sequence SEQ ID NO:25.
 KW parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
 KW recombinant hybrid parvovirus particle;
 KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
 KW Gene; db.
 XX
 OS Adeno-associated virus 6.
 XX
 PN WO2003104392-A2.
 XX
 PD 18-DEC-2003.
 XX
 PF 02-DEC-2002; 2002WO-US038423.
 XX
 PR 18-DEC-2001; 2001US-0341919P.
 XX
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX
 PI Samulski RJ, Rabinowitz JE;
 XX
 DR WPI; 2004-062324/06.
 XX
 XX New polynucleotides comprising parvovirus rep coding sequences and
 PT parvovirus cap coding sequences, useful in producing higher stocks of
 PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
 PT subject.
 XX
 PS Disclosure; SEQ ID NO 25; 115pp; English.
 XX
 CC The present invention describes a polynucleotide (I), comprising
 CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
 CC rep coding sequences encodes a DNA binding domain from a first
 CC parvovirus, and a capsid interacting domain from a parvovirus different
 CC from the first parvovirus. The cap coding sequence comprises sequences
 CC from the different parvovirus. Also described: (1) a vector comprising
 CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
 CC parvovirus cap coding sequences, where the rep coding sequences encode a
 CC DNA binding domain from a first parvovirus and a capsid interacting
 CC domain from a parvovirus different from the first parvovirus, the cap
 CC coding sequences comprise sequences from the different parvovirus, and
 CC the rep coding sequences are stably integrated into the genome of the
 CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
 CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
 CC gene therapy. The polynucleotide (I) can be used in producing higher
 CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in

CC the delivery of nucleic acids having biological effect to treat or
 CC ameliorate the symptoms associated with any disorder related to gene
 CC expression. The polynucleotide may be used to produce a parvovirus vector
 CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
 CC The parvovirus vector may also be used to provide an antisense nucleic
 CC acid to a cell in vitro or in vivo, or in diagnostic and screening
 CC methods. The present sequence is used in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
 Query Match 85.2%; Score 121.8; DB 12; Length 4683;
 Best Local Similarity 98.4%; Pred. No. 1.3e-22;
 Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 19 TTGCCCACTCCTCTCTGCGGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 78
 DB 4683 TTGCCCACTCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 4624
 QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGCGCAGAGAGGAGTG 138
 DB 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGCGCAGAGAGGAGTG 4564
 QY 139 GGCAA 143
 DB 4563 GGCAA 4559
 RESULT 13
 AAD00781/c
 ID AAD00781 standard; DNA; 143 BP.
 XX
 AC AAD00781;
 XX
 DT 08-SEP-2000 (first entry)
 XX
 DE Adeno-associated virus serotype 1 5' inverted terminal repeat.
 XX
 KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
 KW cap protein; recombinant viral vector; gene delivery; gene therapy;
 KW vaccine; transgene; inverted terminal repeat; ITR; ss.
 XX
 OS Adeno associated virus serotype 1.
 XX
 FH Key Location/Qualifiers
 FT misc_binding 1..41
 FT /tag= a
 FT /bound_moiety= "binds to nucleotides 85..125"
 FT stem_loop 42..62
 FT /tag= b
 FT stem_loop 64..84
 FT /tag= c
 FT misc_binding 85..125
 FT /tag= d
 FT /bound_moiety= "binds to nucleotides 41..1"
 FT
 PN WO200028061-A2.
 XX
 PD 18-MAY-2000.
 XX
 PF 02-NOV-1999; 99WO-US025694.
 XX
 PR 05-NOV-1998; 98US-0107114P.
 XX
 PA (UYPE-) UNIV PENNSYLVANIA.
 XX
 PI Wilson JM, Xiao W;
 XX
 DR WPI; 2000-376571/32.
 XX
 XX Novel adeno-associated virus serotype 1 polynucleotide useful for
 PT preparation of medicament for delivery of a transgene to a host.
 PT

Example 2; Fig 2; 108pp; English.

The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions, are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV hence is useful for transforming host cells, and in the preparation of a medicament for the delivery of transgene to a host. The present sequence is a 5' ITR of AAV-1 DNA which is useful in the production of recombinant viral vector. The ITR forms a T-shaped hairpin structure

Sequence 143 BP; 24 A; 50 C; 48 G; 21 T; 0 U; 0 Other;

Query Match 83.2%; Score 119; DB 3; Length 143;
Best Local Similarity 89.5%; Pred. No. 5.9e-22;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
1 TTACCCCTAGTGTGAGTTGCCACTCCCTCTCTCGCGGCTCGCTCGTGGGGCC 60
143 TTACCCCTAGTGTGAGTTGCCACTCCCTCTCTCGCGGCTCGCTCGTGGGGCC 84
61 TCGGACCAAGCTCCGACACCGCAGAGCTCTCTCGCGGCTCGCTCGTGGGGCC 120
83 GGCAGAGCAGAGCTCTCGCGTCTGCGGACCTTTGGTCCGAGCCGCCACCGAGCGCA 24
121 GCGCGCAGAGCGAGTGGGCAA 143
23 GCGCGCAGAGCGAGTGGGCAA 1

RESULT 14

AD000772/c
ID AAD00772 standard; DNA; 4718 BP.
XX
AC AAD00772;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 DNA.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy; vaccine; transgene; ss.
XX
OS Adeno associated virus serotype 1.
XX
FH Key Location/Qualifiers
FT repeat_unit 1..143
FT /tag= a
FT /label= 5' ITR
FT /note= "Inverted terminal repeat which is capable of forming T-shaped hairpin structure"
FT protein_bind 89..110
FT /tag= b
FT /bound moiety= "Rep protein"
FT misc_feature 124..125
FT /tag= c
FT /note= "Terminal resolute site (TRS)"
FT misc_binding 219..226
FT /tag= d
FT /bound moiety= "USF"
FT promoter 236..299
FT /tag= e
FT /label= P5_promoter
FT protein_bind 237..245
FT /tag= f

FT TATA_signal
FT /bound moiety= "YY1 factor"
FT 270..275
FT /tag= g
FT /label= P5_TATA-Box
FT misc_feature 299..306
FT /tag= h
FT /note= "YY1/P5 RNA"
FT CDS 335..2272
FT /tag= j
FT /product= "Rep 68"
FT /function= "regulates replication and integration of AAV DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT CDS 335..2206
FT /tag= i
FT /product= "Rep 78"
FT /function= "regulates replication and integration of AAV DNA into host cell's chromosome"
FT TATA_signal 857..862
FT /tag= l
FT /label= P19_TATA_Box
FT misc_feature 882..883
FT /tag= m
FT /note= "P19 RNA"
FT CDS 1007..2272
FT /tag= o
FT /product= "Rep 40"
FT /function= "regulates replication and integration of AAV DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT CDS 1007..2206
FT /tag= n
FT /product= "Rep 52"
FT /function= "regulates replication and integration of AAV DNA into host cell's chromosome"
FT TATA_signal 1836..1841
FT /tag= p
FT /label= P40_TATA-Box
FT misc_feature 1875..1876
FT /tag= q
FT /note= "P40 RNA"
FT intron 1924..2220
FT /tag= k
FT /note= "This region interrupts the coding sequence of Rep 68 and Rep 40"
FT CDS 2223..4433
FT /tag= r
FT /product= "VP1 protein"
FT /note= "Capsid protein"
FT CDS 2634..4433
FT /tag= s
FT /product= "VP2 protein"
FT /note= "Capsid protein"
FT CDS 2829..4433
FT /tag= t
FT /product= "VP3 protein"
FT /note= "Capsid protein"
FT polyA_signal 4447..4452
FT /tag= u
FT repeat_unit 4576..4718
FT /tag= v
FT /label= 3' ITR
FT /note= "Inverted terminal repeat which is capable of forming T-shaped hairpin structure"
XX
XX WO200028061-A2.
XX
XX 18-MAY-2000.
XX
XX 02-NOV-1999;
XX 05-NOV-1998;
XX 99WO-US025694.
XX 98US-0107114P.
PR

```
XX (UY28-) UNIV PENNSYLVANIA.
XX PA
XX PI Wilson JM, Xiao W;
XX PI
XX WIPI; 2000-376571/32.
XX DR P-P5DB; AAV71161, AAV71164, AAV71165, AAV71166, AAV71167, AAV71168,
XX DR AAY71169.
XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host.
XX
XX Claim 1; Fig 1; 108pp; English.
XX
XX The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
CC characterised by two inverted terminal repeats (ITR) and open reading
CC frames for rep and capsid (cap) proteins. The rep reading frame encodes
CC four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
CC frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
CC sequence or its fragments particularly ITRs, rep and cap coding regions,
CC are useful in production of recombinant viral vectors for gene delivery.
CC These vectors can be used as gene therapy vectors, vaccine vectors or
CC antisense delivery vectors. The AAV-1 does not induce the formation of
CC neutralising antibodies specific to any serotype of AAV hence is useful
CC for transforming host cells, and in the preparation of a medicament for
CC the delivery of transgene to a host
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ
Query Match 83.2%; Score 119; DB 3; Length 4718;
Best Local Similarity 89.5%; Pred. No. 7.2e-22;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGGCGTCTGCTCGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGGCGTCTGCTCGGTGGGCC 84
QY 61 TCGGACCAAGGTCGCGAGAGCGGAGAGCTCTGCTCTGCGGCGCCCAACCGAGCGAGCGA 120
DB 83 GCGAGAGCAGAGCTCTGCGGCTTGGGACCTTTGGTCCGAGGCGCCCAACCGAGCGAGCGA 24
QY 121 GCGCGCAGAGAGGGAGTGGGCAA 143
DB 23 GCGCGCAGAGAGGGAGTGGGCAA 1
RESULT 15
ID ADE76507/c
XX ADE76507 standard; DNA; 4718 BP.
XX AC ADE76507;
XX DT 29-JAN-2004 (first entry)
XX DE Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
XX
XX adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
XX antiarthritic; neuroprotective; antidiabetic; antithyroid;
XX dermatological; antiinflammatory; gene therapy; vaccine;
XX hyperproliferative; cancer; psoriasis; autoimmune disease;
XX rheumatoid arthritis; multiple sclerosis; diabetes;
XX autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
XX Adeno-associated virus 1.
XX OS
XX BP1310571-A2.
XX PN
XX 14-MAY-2003.
XX PD
XX
XX 12-NOV-2002; 2002EP-00257826.
XX PF
XX 13-NOV-2001; 2001US-0350607P.
XX PR
XX 17-DEC-2001; 2001US-0341117P.
XX
```

```
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386675P.
XX
XX (UY28-) UNIV PENNSYLVANIA.
XX PA
XX PI Gao G, Wilson JM, Alvira M;
XX PI
XX WIPI; 2003-450984/43.
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
XX comprises subjecting a sample having a DNA to amplification via
XX polymerase chain reaction.
XX
XX Claim 14; SEQ ID NO 6; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antirheumatic, antithyroid, dermatological, and antiinflammatory. The AAV
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ
Query Match 83.2%; Score 119; DB 10; Length 4718;
Best Local Similarity 89.5%; Pred. No. 7.2e-22;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGGCGTCTGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGGCGTCTGCTCGGTGGGCC 84
QY 61 TCGGACCAAGGTCGCGAGAGCGGAGAGCTCTGCTCTGCGGCGCCCAACCGAGCGAGCGA 120
DB 83 GCGAGAGCAGAGCTCTGCGGCTTGGGACCTTTGGTCCGAGGCGCCCAACCGAGCGAGCGA 24
QY 121 GCGCGCAGAGAGGGAGTGGGCAA 143
DB 23 GCGCGCAGAGAGGGAGTGGGCAA 1
Search completed: November 28, 2005, 19:55:37
Job time : 66.9124 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 522.566 Seconds
(without alignments)
15555.198 Million cell updates/sec

Title: US-10-696-900-1_COPY_4576_4718
Perfect score: 143
Sequence: 1 ttaccctagtgatggagtt.....cgcagagaggagtgggcaa 143

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_ph.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_hcg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	4718	6	BD242766 Adeno-ss
2	143	100.0	4718	6	AR562498 Sequence
3	143	100.0	4718	6	AX753251 Sequence
4	143	100.0	4718	13	AF063497 Adeno-ss
5	135.6	94.8	4721	6	CQ972062 Sequence
6	135.6	94.8	4721	6	CS073592 Sequence
7	135.6	94.8	4721	6	AX753246 Sequence
8	135.6	94.8	4721	13	AF513851 Adeno-ss
9	121.8	85.2	4683	13	AF028704 Adeno-ss
10	120.2	84.1	4683	6	BD242775 Adeno-ss
11	120.2	84.1	4683	6	AR562507 Sequence
12	119	83.2	4718	6	BD242766 Adeno-ss
13	119	83.2	4718	6	AR562498 Sequence
14	119	83.2	4718	6	AX753251 Sequence
15	119	83.2	4718	13	AF063497 Adeno-ss
16	114.8	80.3	4683	13	AF028704 Adeno-ss
17	113.2	79.2	4683	6	BD242775 Adeno-ss
18	113.2	79.2	4683	6	AR562507 Sequence

19	113.2	79.2	4726	6	AX753253 Sequence
20	113.2	79.2	4726	13	AVU48704 Adeno-ss
c 21	111.6	78.0	4721	6	CQ972062 Sequence
c 22	111.6	78.0	4721	6	CS073592 Sequence
c 23	111.6	78.0	4721	6	AX753246 Sequence
c 24	111.6	78.0	4721	13	AF513851 Adeno-ss
25	109	76.2	4681	6	BD242774 Adeno-ss
26	109	76.2	4681	6	AR562506 Sequence
c 27	107.4	75.1	145	6	AR140333 Sequence
28	107.4	75.1	145	6	AX286293 Sequence
29	107.4	75.1	145	13	AA2L7R2 Adeno-ss
30	107.4	75.1	165	6	AR034135 Sequence
31	107.4	75.1	165	6	BD218219 Adeno-ss
32	107.4	75.1	165	6	II6806 Sequence 1
33	107.4	75.1	165	6	AR223306 Sequence
34	107.4	75.1	165	6	AR658389 Sequence
35	107.4	75.1	165	6	AX106702 Sequence
36	107.4	75.1	192	6	A46400 Sequence 3
c 37	107.4	75.1	198	6	AX418199 Sequence
38	107.4	75.1	207	6	AX703496 Sequence
39	107.4	75.1	272	6	BD218218 Adeno-ss
40	107.4	75.1	272	6	AR223305 Sequence
41	107.4	75.1	272	6	AR658388 Sequence
42	107.4	75.1	272	6	AX106701 Sequence
43	107.4	75.1	4675	6	BD094552 Method of
44	107.4	75.1	4675	6	AX135805 Sequence
45	107.4	75.1	4675	6	AX286292 Sequence

ALIGNMENTS

RESULT 1	BD242766	4718 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	Adeno-associated virus serum type 1 nucleic acid sequence, vector				
DEFINITION	Adeno-associated virus serum type 1 nucleic acid sequence, vector				
ACCESSION	BD242766				
VERSION	BD242766.1 GI:33052536				
KEYWORDS	JP 2002529098-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 4718)				
AUTHORS	Wilson,J.M. and Xiao,W.				
TITLE	Adeno-associated virus serum type 1 nucleic acid sequence, vector				
JOURNAL	Patent: JP 2002529098-A 1 10-SEP-2002;				
COMMENT	THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA				
	OS AAV-1				
	PN JP 2002529098-A/1				
	PD 10-SEP-2002				
	PF 02-NOV-1999 JP 2000581227				
	PR 05-NOV-1998 US 60/107114				
	PI JAMES W WILSON,WEIDONG XIAO				
	PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC				
	C12N1/21,				
	PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC				
	Adeno-associated virus serum type 1 nucleic acid sequence, CC				
	vector and host				
	CC cell containing the same				
	PH Key Location/Qualifiers				
	FT CDS (335) . . (2206)				
	FT CDS (2223) . . (4430) .				
FEATURES	Location/Qualifiers				
	source				
	1..4718				
	/organism="unidentified"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32644"				
ORIGIN					
	Query Match 100.0%; Score 143; DB 6; Length 4718;				
	Best Local Similarity 100.0%; Pred. NO. 1.8e-21;				

```
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTATGAGGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 60
Db 4576 TTACCCCTAGTATGAGGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 4635
QY 61 TCGGACCAAAAGTTCGCGAGCGGAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGA 120
Db 4636 TCGGACCAAAAGTTCGCGAGCGGAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGGTGGGCAA 143
Db 4696 GCGCGCAGAGGAGGTGGGCAA 4718

RESULT 2
AR562498
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 4718)
Wilson,J.M. and Xiao,W.
Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
Patent: US 6759237-A 1 06-JUL-2004;
JOURNAL The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
source
1. .4718
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTATGAGGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 60
Db 4576 TTACCCCTAGTATGAGGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 4635
QY 61 TCGGACCAAAAGTTCGCGAGCGGAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGA 120
Db 4636 TCGGACCAAAAGTTCGCGAGCGGAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGGTGGGCAA 143
Db 4696 GCGCGCAGAGGAGGTGGGCAA 4718

RESULT 3
AR573251
LOCUS AR573251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
ACCESSION AR573251
VERSION AR573251.1 GI:32166108
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE
1
Gao,G., Wilson,J.M. and Alvira,M.
A method of detecting and/or identifying adeno-associated virus
(AAV) sequences and isolating novel sequences identified thereby
Patent: EP 1310571-A 6 14-MAY-2003;
JOURNAL The Trustees of The University of Pennsylvania (US)
FEATURES
Location/Qualifiers
source
1. .4718
/organism="Adeno-associated virus 1"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:85106"

ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTATGAGGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 60
Db 4576 TTACCCCTAGTATGAGGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 4635
QY 61 TCGGACCAAAAGTTCGCGAGCGGAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGA 120
Db 4636 TCGGACCAAAAGTTCGCGAGCGGAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGGTGGGCAA 143
Db 4696 GCGCGCAGAGGAGGTGGGCAA 4718

RESULT 4
AF063497
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999
DEFINITION Adeno-associated virus 1, complete genome.
ACCESSION AF063497
VERSION AF063497.1 GI:4689096
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE
1 (bases 1 to 4718)
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and Wilson,J.M.
TITLE Gene therapy vectors based on adeno-associated virus type 1
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)
PUBMED 10196295
REFERENCE
2 (bases 1 to 4718)
Xiao,W. and Wilson,J.M.
Direct Submission
TITLE Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601
JOURNAL Spruce Street, Philadelphia, PA 19104, USA
FEATURES
Location/Qualifiers
source
1. .4718
/organism="Adeno-associated virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:85106"
/note="AAV1"
335. .2206
/codon_start=1
CDS
/product="nonstructural protein"
/protein_id="AAD27758.1"
/db_xref="GI:4689098"
/translaton="MPGYEIVIKVPSDLDEHLPGISDSFVSWVAEKWEWLPDSDMD
LNLIQAPLTVAEKLQDFLVQMRVSKAPALFEVQPKGESYPHLHILVETTVKS
MYLGRFLSQIRDLKVLQTIYRGIEPLPNWFAVTKRNGAGGNGKVVDSYIPNLLPK
TOPELQAWNTMSEYISACLNLAERKRLVAQLTHVSQTQSONKENLNPNSDAPVRS
KTSAPYLVGPAPPADIKTNRIYRIELNGYEPAYAGSVFLGWAQKRFKRNITWLF
PATGTGNTABIAIAHAVPFYGCVMNTNENFPCDVKMVIWMEGKMTAKVVEGAKA
ILGSKVRVDOKSSAOIDPTPTVTSNTNMCAVIDGNSTTFEHOQPLQDRMPKPEL
TRLEHDFGKVTQKVEFFRPAQDHVTEVAHEFYVRKGANKRPAIDDAKSEPKRA
RVSADPSTSDAEGAPVDFADRYQNKSRHAGMLQMLFPFKTCERMNQFNLCPTHGT
RDCSCFFGVGSSEQPVVRKRTYRKLCAIHLHGRAPEIACACDLVNVLDLDCVSEQ"
2223. .4433
/codon_start=1
CDS
/product="capsid protein"
/protein_id="AAD27757.1"
/db_xref="GI:4689097"
/translaton="MAADGYLPDWLEDNLSGIREWDLKPGAPKPKANQKQDDGRG
LVLPGYKYLGPNGLDKGEVNAADAALEHDKAYDQQLKAGDNPYLRNHNADAFQE
RIQEDTSFGNLGRAVFAQKKRVLEPLGLVBEGAKTAPGKKRPVQSQEPDSSSGIQ
KTGQQPAAKRLNFGQTGDSSEVPDPQPLGEPPATPAAVGPTTMSAGGGGAPMADNNEGA
```

DGVGNASGNHCDSTWLGDRVITSTRTRWALFTYNNHLYKQISSASTGASNDNHVFGY
STPWGYFDNRPHCHFSRDRWRLNNWNGFRKLNFKLNIQKVEVITNDGVTTIA
NNLTSVQVFSSEYQLPYVLGSAHQGLPPFADVFMIPQYCYLTLANGSOAVGRSS
FYCLEYFPPSOMLRTGNFTFTSVTFEVPFPHSHYSQSLDRMLNPLIDLYLYLNRTQ
NOSGAONKDLFLPSRPSAGMSVOPKMWLPGPCYRQORVSKTKTNNNSNFTWTGASK
YNLNGRESLINTGTAMASHKDDDEKFFPMSGVMI1GKESAGASNTALDNVMTDSEEL
KATNPVATERFRTGTVAVNFOSSTDPATGDVHWAGALPGMWQORDVYLGOP1WAKLPH
TDGHPHPSLMGGFGLKNPQQLIKNTVPANPABFSAFPIQYTSQGQVSVE
IEWELQKENSKNWNEVQVTSYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL"

ORIGIN

Query Match 100.0%; Score 143; DB 13; Length 4718;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTCGCGCGTCTCGTCTCGTGGGGCC 60
DB 4576 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTCGCGCGTCTCGTCTCGTGGGGCC 4635

QY 61 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGGCCCCACCGAGCGAGCGA 4695

QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGAGGAGTGGGCAA 4718

RESULT 5
CQ972062
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1. .4721
/organism="Adeno-associated virus"
/mol_type="unassigned DNA"
/db_xref="taxon:272636"

ORIGIN

Query Match 94.8%; Score 135.6; DB 6; Length 4721;
Best Local Similarity 97.2%; Pred. No. 7.2e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCT 61
DB 4580 TACCCCTAGTGATGAGTTGCCACTCCCTCTATGCGCGCTCGCTCGTGGGGCCCT 4639

QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 121
DB 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 4699

QY 122 GCGCGCAGAGAGGAGTGGGCAA 143
DB 4700 GCGCGCAGAGAGGAGTGGGCAA 4721

RESULT 6
CQ973592
LOCUS
DEFINITION
ACCESSION

1. .4721
/organism="Adeno-associated virus"
/mol_type="unassigned DNA"
/db_xref="taxon:202812"

ORIGIN

Query Match 94.8%; Score 135.6; DB 6; Length 4721;
Best Local Similarity 97.2%; Pred. No. 7.2e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCT 61
DB 4580 TACCCCTAGTGATGAGTTGCCACTCCCTCTATGCGCGCTCGCTCGTGGGGCCCT 4639

QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 121
DB 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 4699

QY 122 GCGCGCAGAGAGGAGTGGGCAA 143
DB 4700 GCGCGCAGAGAGGAGTGGGCAA 4721

CS073592.1 GI:63090489
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1. .4721
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="adeno-associated virus serotype 7"

ORIGIN

Query Match 94.8%; Score 135.6; DB 6; Length 4721;
Best Local Similarity 97.2%; Pred. No. 7.2e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCT 61
DB 4580 TACCCCTAGTGATGAGTTGCCACTCCCTCTATGCGCGCTCGCTCGTGGGGCCCT 4639

QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 121
DB 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 4699

QY 122 GCGCGCAGAGAGGAGTGGGCAA 143
DB 4700 GCGCGCAGAGAGGAGTGGGCAA 4721

RESULT 7
AX753246
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

1. .4721
/organism="Adeno-associated virus 7"
/mol_type="unassigned DNA"
/db_xref="taxon:202812"

ORIGIN

Query Match 94.8%; Score 135.6; DB 6; Length 4721;
Best Local Similarity 97.2%; Pred. No. 7.2e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCT 61
DB 4580 TACCCCTAGTGATGAGTTGCCACTCCCTCTATGCGCGCTCGCTCGTGGGGCCCT 4639

QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 121
DB 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGCGGCCCCACCGAGCGAGCGAG 4699

QY 122 GCGCGCAGAGAGGAGTGGGCAA 143
DB 4700 GCGCGCAGAGAGGAGTGGGCAA 143

FYCLBYFSPQMLTGNNFTSYTFEDVPHSHSYAHQSQDLRLMNLPLIDQYLYLNRQT
NOSGAQNKLDFSRGSPAGMSVQPKWMLPGCYQRVSKTKTNNNSNFTWTGASK
YNLNGRESTINFGTAMASHKDDKFFPMGSGVMI FKESAGASNTALDNVMTDDEEI
KATNPVATERFCTAVNLQSSSTPATGVDHVMGALPGVMWODRVDYLOQPTWAKIPH
TCHFHPSLKMGGFGLKHPPPOILLKNTFPVPAFPAESATKFAFITYQSTGQSVSE
IEWELQKSKRWNPVEVQTSYNAKSANVDFTVDNNGLYTEPRPIGRYLRPL"

ORIGIN

Query Match 85.2%; Score 121.8; DB 13; Length 4683;
Best Local Similarity 98.4%; Pred. No. 7.1e-17;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 TTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGTGGGCGCTGCGGACAAAGGTCCGC 78
DB 4683 TTGCCACTCCCTCTATGCGCGCTCGCTCGTGGTGGGCGCTGCGGACAAAGGTCCGC 4624

QY 79 AGACGGCAGAGCTCTGCTTGGCGGCCACCGAGCGAGCGGCGGAGGGAGTG 138
DB 4623 AGACGGCAGAGCTCTGCTTGGCGGCCACCGAGCGAGCGGCGGAGGGAGTG 4564

QY 139 GGCAA 143
DB 4563 GGCAA 4559

RESULT 10

BD242775/c
LOCUS BD242775 4683 bp DNA linear PAT 17-JUL-2003
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector
ACCESSION BD242775
VERSION 1 GI:33052545
KEYWORDS JP 2002529098-A/10.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4683)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same
JOURNAL Patent: JP 2002529098-A 10 10-SEP-2002;
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT OS aav-6
PN JP 2002529098-A/10
PD 10-SEP-2002
PF 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES M WILSON,WEIDONG XIAO
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH key Location/Qualifiers
FT source 1. 4683
FT /organism='aav-6'.
Location/Qualifiers
source 1. 4683
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN

Query Match 84.1%; Score 120.2; DB 6; Length 4683;
Best Local Similarity 97.6%; Pred. No. 1.6e-16;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 TTGCCACTCCCTCTGCGCGCTCGCTCGTGGTGGGCGCTGCGGACCAAGGTCCGC 78
DB 4683 TTGCCACTCCCTCTATGCGCGCTCGCTCGTGGTGGGCGCTGCGGACCAAGGTCCGC 4624

QY 79 AGACGGCAGAGCTCTGCTTGGCGGCCACCGAGCGAGCGGCGGAGGGAGTG 138

|||||
DB 4623 AGACGGCAGAGCTCTGCTTGGCGGCCACCGAGCGAGCGGCGGATAGAGGAGTG 4564

QY 139 GGCAA 143
DB 4563 GGCAA 4559

RESULT 11
AR562507/c
LOCUS AR562507 4683 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 19 from patent US 6759237.
ACCESSION AR562507
VERSION AR562507.1 GI:53976573
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4683)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
JOURNAL Patent: US 6759237-A 19 06-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
Location/Qualifiers
source 1. 4683
/organism='unknown'
/mol_type='genomic DNA'

ORIGIN

Query Match 84.1%; Score 120.2; DB 6; Length 4683;
Best Local Similarity 97.6%; Pred. No. 1.6e-16;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 TTGCCACTCCCTCTGCGCGCTCGCTCGTGGTGGGCGCTGCGGACCAAGGTCCGC 78
DB 4683 TTGCCACTCCCTCTATGCGCGCTCGCTCGTGGTGGGCGCTGCGGACCAAGGTCCGC 4624

QY 79 AGACGGCAGAGCTCTGCTTGGCGGCCACCGAGCGAGCGGCGGAGGGAGTG 138
DB 4623 AGACGGCAGAGCTCTGCTTGGCGGCCACCGAGCGAGCGGCGGAGGGAGTG 4564

QY 139 GGCAA 143
DB 4563 GGCAA 4559

RESULT 12

BD242766/c
LOCUS BD242766 4718 bp DNA linear PAT 17-JUL-2003
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same.
ACCESSION BD242766
VERSION BD242766.1 GI:33052536
KEYWORDS JP 2002529098-A/1.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same
JOURNAL Patent: JP 2002529098-A 1 10-SEP-2002;
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT OS AAV-1
PN JP 2002529098-A/1
PD 10-SEP-2002
PF 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES M WILSON,WEIDONG XIAO
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC

PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC

```
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH Key Location/Qualifiers
FT CDS (335)..(2206)
FT CDS (2223)..(4430).
FEATURES
    source
        1..4718
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 2.9e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 84
QY 61 TCGGGACCAAGGTCCGACAGCGGACAGCTTGCTCTGCCCGGCCACCGACGAGCGA 120
DB 83 GGCAGAGCAGAGCTCTGCCGTCTGCGGACCTTTGGTCCGACGCCGCCACCGAGCGA 24
QY 121 GCGCGCAGAGGAGGTGGCAA 143
DB 23 GCGCGCAGAGGAGGTGGCAA 1
RESULT 13
AR562498/c
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
JOURNAL Patent: US 6759237-A 1 08-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
    source
        1..4718
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 2.9e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 84
QY 61 TCGGGACCAAGGTCCGACAGCGGACAGCTTGCTCTGCCCGGCCACCGACGAGCGA 120
DB 83 GGCAGAGCAGAGCTCTGCCGTCTGCGGACCTTTGGTCCGACGCCGCCACCGAGCGA 24
QY 121 GCGCGCAGAGGAGGTGGCAA 143
DB 23 GCGCGCAGAGGAGGTGGCAA 1
RESULT 14
AX753251/c
LOCUS AX753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
```

```
AX753251
VERSION AX753251.1 GI:32166108
KEYWORDS Adeno-associated virus 1
SOURCE Adeno-associated virus 1
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1
AUTHORS Gao,G., Wilson,J.M. and Alvira,M.
TITLE A method of detecting and/or identifying adeno-associated virus
(AVV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
    Location/Qualifiers
        source
            1..4718
                /organism="Adeno-associated virus 1"
                /mol_type="unassigned DNA"
                /db_xref="taxon:85106"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 2.9e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 84
QY 61 TCGGGACCAAGGTCCGACAGCGGACAGCTTGCTCTGCCCGGCCACCGACGAGCGA 120
DB 83 GGCAGAGCAGAGCTCTGCCGTCTGCGGACCTTTGGTCCGACGCCGCCACCGAGCGA 24
QY 121 GCGCGCAGAGGAGGTGGCAA 143
DB 23 GCGCGCAGAGGAGGTGGCAA 1
RESULT 15
AF063497/c
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999
DEFINITION Adeno-associated virus 1, complete genome.
ACCESSION AF063497
VERSION AF063497.1 GI:4689096
KEYWORDS Adeno-associated virus 1
SOURCE Adeno-associated virus 1
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and
Wilson,J.M.
TITLE Gene therapy vectors based on adeno-associated virus type 1
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)
PUBMED 10196295
REFERENCE 2 (bases 1 to 4718)
AUTHORS Xiao,W. and Wilson,J.M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601
Spruce Street, Philadelphia, PA 19104, USA
FEATURES
    Location/Qualifiers
        1..4718
            /organism="Adeno-associated virus 1"
            /mol_type="genomic DNA"
            /db_xref="taxon:85106"
            /note="AAV1"
            335..2206
                /codon_start=1
                /product="nonstructural protein"
                /protein_id="AAD2758.1"
                /db_xref="GI:4689098"
                /translation="MPGFYEVIVKVPISDLDEHLPGISDFSVSWAEKEWELPPDSMD
                LNLIEQALPTVAELKQDFLVQMRVRSKAPALFFVQPEKGESYFHLHLIVETTVKKS
                MYLGRFLSQIRDKLVQTIYRGTEPLPNWFAVTRKNGAGGNGKVVDCYIIPNLLPK
                TOPELQWANTMMEYIISACLNLAEKRLVAOHLTHVSQTQEQNKENLPNSDAPVIRS
                KTSARYMELVGLVDRGITSEKQWIQEDQASYISFNAAASNSRSQIKAAALDNAGKTMAL"
CDS
```

TKSAPDYLGPAPPADIKTNRIYRIELENGYEPAYAGSVFLGWAQKRFGRNTIWLFG
PATTKNTIAEAIAHVPFYGCNVNTNENFPFNDVCVRNVIWHEEGMTAKVVEGAKA
ILGGSXRVQDQKCSAQIDPTPIVITSNTNMCVIDGNSITTFHQOPLQDRMFKEFEL
TRLEHDFGKVKQEFERWAQHVTEVAHEFYVRKGANKRPAPDDADKSEPKRA
CPSVADPSTDAEGAPVDPADRYQNKCSRHAGMLQWLFCKTCERNQNFICFTHGT
RGCSECFPGVSESQPVVRKRYRKLCIAIHLLGRAPETIACSACDLVVDLDDCVSEQ"
2223. .4433
/codon_start=1
/product="capsid protein"
/protein_id="AAD27757.1"
/db_xref="GI:4689097"
/translations="MAADGYLPDWLEDNLSEGREWWDLPKGAPKPKANQKQDDGRG
LVLPGYKILGPNGLDKGEPFNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE
RLQEDTSFGNLRGAVFQAKRVLEPLGEBGAKTAPGKRPEQSPQEPDSSSGIG
KTGQQAQKRLNFGQTSVDPQPLGEPFATPAAGVPTTMASSGAGPMDNNEGA
DGVNAGNWHCDSTWLGDRVITTSRTWALPTYNNHLYKQISSASTGASNDNHFGY
STPMGYFDNBRHCHESPRDMORLINNHGFRPKRLNFKLENIQVKEVTNDQVTTIA
NMLTSTVQFSSEYQLPYVLGSAHQCLPPPADVPMIPQYGYLTLNNGSQAVERSS
FYCLEYFPQMLRTGNFTFSYTFEEVFPFSSYAHQSILDRMLNPLIDQLLYLNLRTQ
NOSGAQNKDLILFSRGSAGMSVQPKWLPFGPCYQQRVSKTKTDNNNSNFTWTGASK
YNLNGRESIINPGTAMASHKDDKDFPMSGVMI FGKESAGASNTALDNVMTDEBEI
KATNPVATERFGTVAVNFOSSTDPATGDVHAMGLPGMWQDRDVLQGPWAKIPH
TDGHFPLMCGFGLKNPPQILIKNTVPANPPAEFSATKFAFITQYSTGOVSVE
IEWELQKNSKRWNPVQYTSNTAKSANVDFTVDNNGLYTEPRPIGTRILTRPL"

ORIGIN

Query Match 83.2%; Score 119; DB 13; Length 4718;
Best Local Similarity 89.5%; Pred. No. 2.9e-16;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 1 TTACCCCTAGTGATGGAGTTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
Db 143 TTAGCCCTAGTGATGGAGTTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 84
Qy 61 TCGGACCAAGCTCCGACACGCGAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGA 120
Db 83 GGCAGAGCAGAGCTCTGCGGACCTTTGGTCCCGAGGCCCCACCGAGCGAGCGA 24
Qy 121 GCGCGCAGAGGGAGTGGGCAA 143
Db 23 GCGCGCAGAGGGAGTGGGCAA 1

Search completed: November 29, 2005, 00:05:26
Job time : 523.566 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 454.049 Seconds
(without alignments)
14735.325 Million cell updates/sec

Title: US-10-696-900-1_COPY_4576_4718
Perfect score: 143
Sequence: 1 ttaccctagtgatggagt.....cgcagagggagtgaggcaa 143

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_hcc:*
 - 5: gb_est4:*
 - 6: gb_est5:*
 - 7: gb_est6:*
 - 8: gb_est7:*
 - 9: gb_gsl:*
 - 10: gb_gsl2:*
 - 11: gb_gsl3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.8	26.4	763	2	BG490707 602520211
C 2	37.8	26.4	1021	5	B0838990 AGENCOURT
C 3	37.4	26.2	644	2	BB632576 BB632576
C 4	36.2	25.3	604	3	BI531460 1024113H0
C 5	36.2	25.3	932	10	AL066742 Drosophila
C 6	36.2	25.3	1021	5	B0838990 AGENCOURT
C 7	36	25.2	318	7	CV352119 MR3-GN018
C 8	35.8	25.0	1074	10	CL507458 SAIL_781
C 9	35.6	24.9	578	2	BB651469 BB651469
C 10	35.6	24.9	793	3	BI905802 603062735
C 11	35.6	24.9	1115	3	BM906771 AGENCOURT
C 12	35.6	24.8	1123	10	AG080476 Pan trogl
C 13	35.2	24.6	2024	10	AG074763 Pan trogl
C 14	35	24.5	724	10	CL751558 OR_Ba011
C 15	35	24.5	844	10	AL056652 Drosophila
C 16	34.8	24.3	676	9	BZ414423 if25a12.9
C 17	34.6	24.2	617	7	CN038418 nm.27.k20
C 18	34.6	24.2	637	7	CN045740 vll_p57.1
C 19	34.6	24.2	695	10	CZ814168 OC_Ba019
C 20	34.6	24.2	860	2	BF578052 602094705
C 21	34.6	24.2	926	8	CV775269 FGAS06967
C 22	34.6	24.2	2088	10	CL507788 SAIL_788

C 23	34.4	24.1	671	10	CW914921
C 24	34.4	24.1	890	10	AG127787
C 25	34.4	24.1	1136	3	BM416106
C 26	34.2	23.9	364	5	BY014749
C 27	34.2	23.9	468	8	CX358645
C 28	34.2	23.9	914	3	BI762701
C 29	34.2	23.9	1009	10	CNS010EW
C 30	34	23.8	343	5	BY181775
C 31	34	23.8	450	5	BY250974
C 32	34	23.8	623	2	BB613613
C 33	34	23.8	631	5	BY723537
C 34	34	23.8	639	2	BB651977
C 35	34	23.8	643	2	BB652152
C 36	34	23.8	657	2	BB651054
C 37	34	23.8	661	2	BB621081
C 38	34	23.8	678	5	BY726305
C 39	34	23.8	795	9	AQ331299
C 40	34	23.8	838	8	CV780024
C 41	34	23.8	1134	6	CA459340
C 42	34	23.8	1364	5	BU543298
C 43	34	23.8	1767	4	AK082810
C 44	34	23.8	2268	4	AK082478
C 45	34	23.8	2708	4	AK038836

ALIGNMENTS

RESULT 1
BG490707/c
LOCUS
DEFINITION
602520211F1 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:4638699 5',
mRNA sequence.
ACCESSION
BG490707
VERSION
BG490707.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 763)
NIH-MGC http://mgc.mci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DP/Genetics
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1401 row: e column: 04
High quality sequence stop: 182.

FEATURES
source
1..763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4638699"
/tissue_type="large cell carcinoma"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH_MGC_18"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 26.4%; Score 37.8; DB 2; Length 763;
 Best Local Similarity 58.4%; Pred. No. 13;
 Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 26 CTCCTCTCTGCGGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGACAGCGC 85
 DB 261 CGCTTGGCTGCGGGGCGCTCGCTCTGCTCTGCCCGCCCTCGCTTCGCTCGGC 202

QY 86 AGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCAGAGGAGGAGTG 138
 DB 201 AGCGGTACGCTTGGCGGCGCTGACGCGGAGGAGCGGCTGTCGCGGCGTG 149

RESULT 2

LOCUS BU838990/c 1021 bp mRNA linear EST 16-OCT-2002
 DEFINITION AGENCOURT 8209898 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6258289
 5' mRNA sequence.

ACCESSION BU838990
 VERSION BU838990.1 GI:24023385
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominiidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-remail.nih.gov

Tissue Procurement: DCTP/DRP

cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LLCM2413 row: p column: 02

High quality sequence start: 36

High quality sequence stop: 231.

FEATURES

source

1..1021
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6258289"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 112"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCACAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Query Match 26.4%; Score 37.8; DB 5; Length 1021;
 Best Local Similarity 58.4%; Pred. No. 13;
 Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 28 CCCTCTCTGCGGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGACAGCGCAG 87
 DB 903 CGGCGCGCGGCGCTAGCGGGGCGGGGCGCCCTCCGCGCGCGTCCGGGGGCGCGG 844

QY 88 AGCTCTGCTCTGCGGCGCCCAACCGAGCGAGCGAGCGCGCAGAGAGGAGTG 140

Db

843 CGACGGGGGAAGCGCGCCACCGCGGCGCTCGCGGGGTGGTGGGG 791

RESULT 3

LOCUS BB632576

DEFINITION

BB632576 RIKEN full-length enriched, adult male hypothalamus Mus
 musculus cDNA clone A230067E15 5', mRNA sequence.

ACCESSION

BB632576

VERSION

BB632576.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 644)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,
 Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,
 Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,
 Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 Unpublished (2001)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-3222
 Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
 and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,

Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and

Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

e mouse tissues.

Location/Qualifiers

1..644

FEATURES

source

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="A230067E15"

/sex="male"

/tissue_type="hypothalamus"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, adult male

hypothalamus"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Db 498 CGACAAAGCGCCCTCCGCTGGCGCGCTACGGGGAGCGCGGCTCAATTACGAGGAG 557

Qy 139 GGC 141

Db 558 GGC 560

RESULT 15	CNS0052P	844 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
DEFINITION	BACR11P6 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL056652				
VERSION	AL056652.1	GI:4932342			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 844)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequences : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqres@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
SOURCE	1..844 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACR11P6" /clone_lib="RPCI-98" /notes="Tend : TET3"				

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 5092.96 Seconds
(without alignments)
14735.325 Million cell updates/sec

Title: US-10-696-900-1_COPY_2829_4432
Perfect score: 1604
Sequence: 1 atgggttcaggcggtggcgc.....tacctaccgctccctgtga 1604

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	180.8	11.3	264	7	CO892248
C 2	126.8	7.9	195	7	CO888893
C 3	66.8	4.2	753	9	BH115587
C 4	48	3.0	367	3	BJ432209
C 5	47	2.9	581	5	BW335572
C 6	45.6	2.8	681	10	CNS02EOD
C 7	45.6	2.8	869	7	CK159167
C 8	45.4	2.8	581	6	CF569196
C 9	43.6	2.7	414	1	AV964990
C 10	43.4	2.7	801	7	CO465395
C 11	43.2	2.7	748	7	CO048199
C 12	42.8	2.7	415	2	BE556771
C 13	42.8	2.7	668	6	CD920707
C 14	42.2	2.6	435	5	BQ449242
C 15	42.2	2.6	450	3	BJ371297
C 16	42.2	2.6	473	2	BE016852
C 17	42.2	2.6	477	6	CA743784
C 18	42.2	2.6	530	1	AW566595
C 19	42.2	2.6	546	5	BW249566
C 20	42.2	2.6	552	3	BQ285353
C 21	42.2	2.6	660	5	BW193819
C 22	41.8	2.6	539	7	CR929543

23	41.8	2.6	939	10	CNS00CNG
24	41.6	2.6	536	3	BI774214
25	41.6	2.6	599	7	CN788278
26	41.6	2.6	655	7	CK831007
27	41.6	2.6	717	7	CK982139
28	41.4	2.6	620	8	CX180847
29	41.4	2.6	658	5	BX622561
30	41.4	2.6	1116	7	CN068332
31	41.2	2.6	359	3	BI881431
32	41.2	2.6	510	3	BM572584
33	41.2	2.6	546	9	BH589499
34	41.2	2.6	549	3	BM859438
35	41.2	2.6	675	5	BM520599
36	41.2	2.6	713	5	BW534121
C 37	41	2.6	476	1	AL720322
38	41	2.6	612	2	BE195101
39	41	2.6	670	6	CB047069
40	41	2.6	690	7	CK661380
41	41	2.6	720	6	CA296025
42	41	2.6	4479	4	BC044003
43	40.8	2.5	644	3	BM605179
44	40.6	2.5	671	5	BQ987654
45	40.6	2.5	690	5	BQ988583

ALIGNMENTS

RESULT 1
CO892248/c
LOCUS
DEFINITION
CO892248
CO892248.1 GI:51822548
EST.
SOURCE
ORGANISM
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 264)
Hennig S., Janitz M., Herwig R. and Williams J.
Generation, annotation, evolutionary analysis and database
integration of 14969 cattle EST clusters
Unpublished (2004)
Contact: Hennig S
Laboratory 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inhestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennigmolgen.mpg.de
The library was characterized by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones that display the same hybridization matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONFP cluster was selected for sequencing. cDNA clones and
filters are distributed via Deutsches Ressourcenzentrum fuer
Genomforschung GmbH (<http://www.rzpd.de>).
PCR Primers
FORWARD: 5' GCTATTACCGACTTGTATGCTTCGGGTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GGTATTACCGACTTGTATGCTTCGGGTCG 3' (M13FSP) 3'-seq
Seq primer: 5'-CCGTCGGGATTCGGGT-3' (M13RSP).
Location/Qualifiers
1. .264
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="RZPDp1056M0360Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"

AL059400 Drosophil
BI774214 466238 MA
CN788278 4122596 B
CK831007 4054395 B
CK982139 4114934 B
CX180847 B04_45-77
BX622561 BX622561
CN068332 Mill_Ag2_P
BI881431 fm92c02.Y
BM572584 fx55a08.Y
BH589499 BGIU064TF
BM859438 fy56a04.Y
BW520599 BM520599
BW534121 BW534121
AL720322 AL720322
BE195101 HVM5H008
CB047069 NISC_gf08
CK661380 LP20345.5
CA296025 SCAGLV104
BC044003 Xenopus 1
BM605179 170006870
BQ987654 QGF12N04.
BQ988583 QGF15E24.

/clone lib="normal cattle brain"
/note="Organ: brain; Vector: pSport1; Site 1: NotI;
Site 2: SalI; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGACTAGTTCTAGATCGGAGCGCGGCC (T)15-3' and SalI 5'-
TCGACCCACGCGTCCG-3' adapters (Gibco BRL))"

ORIGIN

Query Match 11.3%; Score 180.8; DB 7; Length 264;
Best Local Similarity 80.3%; Pred. No. 5.2e-42;
Matches 212; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1241 GTCCCATTTGGGCCAAATTCCTCACACAGATGGACATTTCCACCGTCTCTCTATGG 1300
DB |||||
264 GTCCCATTTGGGAATGATTCACACTCGGAGCGACATTTTCCACCCCTCTCCCTCATGG 205
DB |||||
QY 1301 GCGGCTTTGGACTCAAGAACCCGCTCTCAGATCCTCATCAAAACACGCGCTTTCCTG 1360
DB |||||
204 GGGGATTTGGACTTAATCAACCCTCTCCACAGATTTTCATCAAGAACACGCGGTTCTCTG 145
DB |||||
QY 1361 CGAATCTCCGGGGAGTTTTCAGCTACAAAGTTTGTCTTCATTCATCACCCTAATCTCCA 1420
DB |||||
144 CGAGTCTCTCGAGCAGCTTCAGTGGCGGAAAGTTTGTCTTCATTCATCACAGTACTCCA 85
DB |||||
QY 1421 CAGGACAAGTGTGTTGAAATTTGAATGGAGCTGCAGAAAGAAACACAGACGCTGGA 1480
DB |||||
84 CGGTCAGTCAAGTGGAGATTTAGTGGAGCTGGGAGGAAACACAGAAACGCTGGA 25
DB |||||
QY 1481 ATCCCGAAGTGCAGTACATCCA 1504
DB |||||
24 ATCCCGAATTCAGTACACTCCA 1

RESULT 2

CO888893/c
LOCUS BovGen17218 normal cattle brain Bos taurus cDNA clone EST 01-SRP-2004
DEFINITION RZPDp105600960Q 5', mRNA sequence.

ACCESSION CO888893
VERSION CO888893.1 GI:51819178
KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.

REFERENCE Hennig S., Janitz M., Herwig R. and Williams J.

AUTHORS Generation, annotation, evolutionary analysis and database

TITLE Integration of 14969 cattle EST clusters

JOURNAL Unpublished (2004)

COMMENT

Contact: Hennig S
Laboratory 123, Dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr. 63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONFP cluster was selected for sequencing. cDNA clones and
filters are distributed via Deutsches Ressourcenzentrum fuer
Genomforschung GmbH (<http://www.rzpd.de>).

PCR Primers

FORWARD: 5' CCCAGCTTTACATTTAGTTCGGCTCG 3' (M13RP) 5'-seq

BACKWARD: 5' GCTATACCGAGCTGGGAAAGGGGATGTG 3' (M13RP) 3'-seq

Seq primer: 5'-CCGTCGGAATTCGCGGT-3' (M13RP).

FEATURES

source

Location/Qualifiers

1..195

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"
/clone="RZPDp105600960Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"
/clone lib="normal cattle brain"
/note="Organ: brain; Vector: pSport1; Site 1: NotI;
Site 2: SalI; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGACTAGTTCTAGATCGGAGCGCGGCC (T)15-3' and SalI 5'-
TCGACCCACGCGTCCG-3' adapters (Gibco BRL))"

ORIGIN

Query Match 7.9%; Score 126.8; DB 7; Length 195;
Best Local Similarity 78.4%; Pred. No. 5.8e-26;
Matches 152; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1411 CAATACTCCACAGACAAAGTGTGGAAATTTGAATGGAGCTGCAGAAAGAAAACAGC 1470
DB |||||
194 CAGTACTCCACGGACAGGTGAGTGGAGATCGATGGGAGTTGCAGAGGTAACAGC 135
DB |||||
QY 1471 AAGCGCTGGAATCCGAAGTGCAGTACATCCAAATTTATGCAAAATCTGCCAAGTTGAT 1530
DB |||||
134 AAACGGTGGAAATCCGAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGAT 75
DB |||||
QY 1531 TTTACTGTGGACAACATGGACTTTTACTGAGCTCGCCCATTTGGCACCGTTACCTT 1590
DB |||||
74 TTTAATGTGGACATTTAATGGCGTGTATTTCAGAGCTCGCCCATTTGGCACCATACCTG 15
DB |||||
QY 1591 ACCCGTCCCTGTGA 1604
DB |||||
14 ACTGTAATGTGTA 1

RESULT 3

BH115587

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Sciurognathi; Muridea; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Tsagaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Akimret, B., Levins, M.,

Russell, D., de Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-24

Unpublished (1999)

Other GSRS: RPCI-24-358F16.TJ

CONTACT: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: azhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC

library availability, please contact Pieter de Jong

(pdjong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end

plate: http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html

Seq primer: 17

Class: BAC ends.

Location/Qualifiers

1..753

/organism="Mus musculus"

/mol_type="genomic DNA"

```
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPC1-24-358F16"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPC1-24"
/Note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;
The RPC1-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
ORIGIN
Query Match 4.2%; Score 66.8; DB 9; Length 753;
Best Local Similarity 64.6%; Pred. No. 7.6e-08;
Matches 115; Conservative 0; Mismatches 62; Indels 1; Gaps 1;
QY 1 ATGGCTTACGGGTGGCGCACCACCAATGGCAGACAAATACGAAGGCGCCGACGAGTGGGT 60
DB 521 ATCTCTGGAGCTGGAGGTCTACTAATGGGCAACAACAC-CAGACACTGATGGAGTGAGC 579
QY 61 AATGCTCTAGGAATATGGGATTCGATTCACATGCTGGCGCAGAGTCTATCACCACC 120
DB 580 AATGCCATGGATTTGGCATTCGCCATTCCAGTGTATGTAGTCTAGTCTATTAATCTAA 639
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACATTAACCACTCTTACAAAGCAAAATCTCCA 178
DB 640 TCTCCCGACGACCTGGGCTCTGCCCACTCAACTAACATTTCCACAACATATATGAACA 697
RESULT 4
BJ432209/c 367 bp mRNA linear EST 13-MAR-2002
LOCUS BJ432209 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION dictyostelium cDNA clone ddv18402 3', mRNA sequence.
ACCESSION BJ432209
VERSION BJ432209.1 GI:19406931
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 367)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source 1..367
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv18402"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostelium discoideum cDNA library, VF"
ORIGIN
Query Match 3.0%; Score 48; DB 3; Length 367;
Best Local Similarity 49.6%; Pred. No. 0.023;
Matches 120; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
QY 110 TCATCACCACGAGCACCCTGGGCGCTTGGCCCTTACCAATACCACTCTACAGC 169
DB 328 TCAACAACATCANCAACATCAATCTCTAAACCAACCAACTATCAACCAACATCAACT 269
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPC1-24-358F16"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPC1-24"
/Note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;
The RPC1-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
ORIGIN
Query Match 2.9%; Score 47; DB 5; Length 581;
Best Local Similarity 45.9%; Pred. No. 0.055;
Matches 161; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
QY 83 GCGATTCCACATGGCTGGGCGCAGAGTCTATCACCACCGACCCCGCCTGGCTTGC 142
DB 148 GCNACTAGAGACGACCAACACATCATCAACACACGAGCGGCGCAACAGCAACA 207
QY 143 CCACCTACAAATACCACTCTTCAAGCAAAATCTCCAGTGTCTTCAACGGGGGCGAGCAG 202
DB 208 CCAACAACCAACCAACAACAACCAACAGCAGCAACAGGAACAACAATCACCACA 267
QY 203 AGAACCACTACTTGGGCTACGACCCCTGGGGGTATTTTGATTTTCAACAGATTCCACT 262
DB 268 ACAACAGCATCATCAACAACAACAGCATCATCCAAACAACAGGAACAACAACAGCACA 327
QY 263 GCCACTTTTCCACCATGCTGGCAGGACTCATCAACAACAAATTTGGGATTTCCGGCCCA 322
DB 328 GGAACAACAACGCGCACAACAGGAACAACAACAGGAACAACAACGCGCACCACCAAGGA 387
QY 323 AGAGACTCAACTTCAAACTCTTCAACATCCAGTCAAGGAGGTCAACGACGAATGATGGCG 382
```

```

Db      388  ACAACAGGAAACAACACAGCACCACAGGAAACAACAGCAGCAACAGCAATCATCAATA 447
      383  TCACACACCTCGCTAATACCTTACACAGCAGCGTTCAGTCTTCGCGACT 433
      448  ACAACACACGCTAATACAAGAAACAACACAAACACATCACCACCACT 498

RESULT 6
CNS02EOD/LOCUS      681 bp  DNA  linear  GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
              262H14 of library G from Tetraodon nigroviridis, genomic survey
              sequence.
ACCESSION  AL193990
VERSION    AL193990.1 GI:7832096
KEYWORDS  GSS; genome survey sequence.
SOURCE    Tetraodon nigroviridis
ORGANISM  Tetraodon nigroviridis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
           Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
           Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE  1
AUTHORS   Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
           Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
           Saurin,W. and Weissenbach,J.
           Estimate of human gene number provided by genome-wide analysis
           using Tetraodon nigroviridis DNA sequence
           Nat. Genet. 25 (2), 235-238 (2000)
           10835645
REFERENCE  2
AUTHORS   Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
           Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
           Saurin,W., Bernot,A. and Weissenbach,J.
           Characterization and repeat analysis of the compact genome of the
           freshwater pufferfish Tetraodon nigroviridis
           Genome Res. 10 (7), 939-949 (2000)
           10899143
REFERENCE  3 (bases 1 to 681)
AUTHORS   Genoscope.
           Direct Submission
           Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]
           Web : www.genoscope.cns.fr
COMMENT   This sequence is a single read and was generated as part of a large
           scale clone-end sequencing project of the Tetraodon nigroviridis
           genome. For more information, please take a look at
           http://www.genoscope.cns.fr/Tetraodon.
FEATURES  Location/Qualifiers
           1..681
           /organism="Tetraodon nigroviridis"
           /mol_type="genomic DNA"
           /db_xref="taxon:99883"
           /clone="262H14"
           /clone_lib="G"
           /note="Genoscope sequence ID : C0AG262DD07LP1
           end : T7"
ORIGIN
Query Match      2.8%; Score 45.6; DB 10; Length 681;
Best Local Similarity 43.4%; Pred. No. 0.15;
Matches 144; Conservative 10; Mismatches 178; Indels 0; Gaps 0;

QY      31  GACAATAACGAGCGCCGACGAGTGGTAAATGCTCAGGAAATTTGGCATTTGCCATTC 90
      351  GACAACAACGACACACACGACACACAAACGACGCGCGACGACGACGACAC 292

QY      91  ACATGGCTGGGCGACAGAGTCATCACCACGACCGCAGCTGGCCCTTGGCCACTTAC 150
      291  RACGGCAGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 232

QY      151  AATAACCACTCTACAAGCAAAATCTCCAGTGTCTTCAACGGGGGCCGACGACCAAC 210

```

```

Db      231  AACACACACACACACAAACAAACACACACACACACACACACACACACACACACAC 172
      211  TACTTTGGCTACAGCACCCCTGGGGGTATTTTGTATTTCAACAGATTTCACCTGCTTT 270
      171  AAAAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 112
      271  TCACACGCTGACTGCAGGAGCTCATCAACAACAATTGGGGATTTCGGGCCCAAGAGACTC 330
      111  AACAAAAACACACACACACACACACACACACACACACACACACACACACACACACAC 52
      331  AACTTCAAACTCTTCAACATCAAGTCAAGGA 362
      51  AASMACACACACACACACACACACACACACACACACACACACACACACACACACAC 20

CKL59167      869 bp  mRNA  linear  EST 05-DEC-2003
FGAS040564 Tetraodon aestivum FGAS: Talt5 Tetraodon aestivum cDNA,
mRNA sequence.
CKL59167
CKL59167.1 GI:38985053
EST.
Tetraodon aestivum (bread wheat)
Tetraodon aestivum
Tetraodon aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE  1 (bases 1 to 869)
AUTHORS   Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
           Genewein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
           Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,
           Pennik,C., Roach,J.L. and Sarhan,F.
           Functional Genomics of Abiotic Stress in Wheat and Canola Crops
           Unpublished (2003)
           Contact: Wm L Crosby
           Bioinformatics
           University of Saskatchewan, Department of Computer Science
           1C101 Engineering Building, 57 Campus Drive, Saskatoon,
           Saskatchewan, S7N 5A9, Canada
           Tel: 306 966 1769
           Fax: 306 966 2033
           Email: fgas_estecs.usask.ca
COMMENT   This sequence is the direct result of the Base calling software
           Phred (default parameters). It is the raw base calls. To aid in the
           identification of the high quality insert the software Lucy
           (default parameters) has been run on this sequence. Lucy identified
           the region [128,636].
           Plate: Talt537 row: N column: 23.
FEATURES  Location/Qualifiers
           1..869
           /organism="Triticum aestivum"
           /mol_type="mRNA"
           /cultivar="Wheat line PI 178383"
           /db_xref="taxon:4565"
           /lab_host="DH5 alpha"
           /clone_lib="Triticum aestivum FGAS: Talt5"
           /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
           subtractive hybridization) cDNA library from genotype
           PI178383 cold hardened at 2 C for 21 days and 49 days
           (equal amount of cDNA pooled together before subtraction,
           tester) and subtracted against genotype Norstar cold
           hardened at 2 C for 1 day (24 H)(driver). Modified Smart
           cDNA (Clontech)priming and non-directional cloning"
ORIGIN
Query Match      2.8%; Score 45.6; DB 7; Length 869;
Best Local Similarity 48.1%; Pred. No. 0.17;
Matches 129; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY      145  ACCTACAATAACCACTCTACAAGCAAAATCTCCAGTGTCTTCAACGGGGGCCGACGAC 204

```

[illegible]

RESULT 8	CF569196	581 bp	linear	EST 08-SEP-2004
LOCUS	CF569196			
DEFINITION	EST057 Subtracted, Clontech (cat. # K1804-1) <i>Triticum aestivum</i> CDNA			
ACCESSION	CF569196			
VERSION	CF569196.1	GI:51921536		
KEYWORDS	EST.			
SOURCE	<i>Triticum aestivum</i> (bread wheat)			
ORGANISM	<i>Triticum aestivum</i>			
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.			
AUTHORS	Xiao,K., Bai,G.H. and Carver,B.F.			
TITLE	Nylon Filter Arrays Reveal Differential Expression of Expressed Sequence Tags in Wheat Roots Under Aluminum Stress			
JOURNAL	J. Integr. Plant Biol. 47 (7), 839-848 (2005)			
COMMENT	Contact: Guihua Bai USDA/ARS and Department of Agronomy Kansas State University Manhattan, KS 66506, USA Email: ghai@bear.agron.ksu.edu Seq primer: M13 Forward High quality sequence stop: 581. Location/Qualifiers 1. .581			
FEATURES				

```

location/Qualifiers
1. .581
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="OK91G106"
/db_xref="taxon:4565"
/clone="FDC57"
/tissue_type="root"
/clone_lib="Subtracted, Clontech (cat. # K1804-1)"
/notes="EST from wheat (Triticum aestivum, cv. OK91G106)
root in response to aluminum stress"

```

Query Match	2.8%	Score 45.4;	DB 6;	Length 581;
Best Local Similarity	46.6%;	Pred. NO. 0.16;		
Matches 145;	Conservative	0;	Mismatches 166;	Indels 0;
Gaps 0;				

Qy	102	CGACAGAGTCAATACCAACGACCCCGCACTTGGGCTTGCCCACTTCAATATACCACCT	161
Db	152	CAACAACAACAACACCAACCAACAGCAGCAACAGCAACAAACAGCAACAACAA	211
Qy	162	CTACAAGCAATCTTCAGTGTCTTCAACGGGGGCGAGCAACGACACCACTACTTCGGCTA	221
Db	212	CAACAACAACAACAGCAACAACAACAGCAACAACAACAACATCAGCAACAACAA	271
Qy	222	CAGCACCCCTGGGGGTATTTTGATTTTCAAAGATTTCCACTTGCACCTTTTTCACACGTGA	281
Db	272	CAGCAGACCAACACAGCAGCAACAACAGCAGCAGCAGCAACAGCAACAGTATCAACAGCA	331

Qy	282	CTGGCAGCAGCTCATCAACAACTTGGGGATTCGGCCCAAGAGACTCACTTCAACT	341
Db	332	CAACAGCAGCAACACCAACAAACAACAACAACAACAACAACAACAACAACA	391
Qy	342	CTTCAACATCCAAGTCAGGAGGTCAGGAGCATGATGGCGTCACACCACTCGCTAATAA	401
Db	392	CAGCAACAACAACAACAACAACATCAGCAACAACATCAGCAACAACAGCAGCAGCA	451
Qy	402	CCTTACCAGCA	412
Db	452	CAGCAACAGCA	462

RESULT 9	AV964990/c	LOCUS	AV964990	414 bp	mRNA	linear	EST 14-MAR-2002
DEFINITION			AV964990	Nozi Satoh unpublished cDNA library, young adult Ciona intestinalis cDNA clone ciad17116 5'.		mRNA sequence	

ACCESSION	AV964990.1	GI:19454686
VERSION	AV964990.1	GI:19454686
KEYWORDS	EST.	
SOURCE	Ciona intestinalis	
ORGANISM	Ciona intestinalis	
REFERENCE	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.	
AUTHORS	1 (bases 1 to 414)	
TITLE	Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.	
JOURNAL	Expressed genes in Ciona intestinalis	
COMMENT	Unpublished (2000)	
	Contact: Nori Satoh	

Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

```

FEATURES
source
Location/Qualifiers
i. .414
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciadi7116"
/tissue_type="whole animal"
/dev_stage="young adult"
/clone_lib="Nori Satoh unpublished cDNA library, young
adult"

```

ORIGIN

Query Match	2.7%	Score 43.6;	DB 1;	Length 414;
Best Local Similarity	46.6%;	Pred. No. 0.49;		
Matches 139;	Conservative	0;	Mismatches 159;	Indels 0;
Gaps 0;				

Qy	111	CATCACACAGACCGCAGACTTGGGGCCTTGGCCACCTACAAATTAACAACCTTCTACAGCA	170
Db	320	CAACAAACAACAACCTGTCTTACAGGCACAACAACAACACCTGTATTCAACAACAACA	261
Qy	171	AATCTCAGATGCTTCAACGGGGGCAGCAACGACAACACTACTTTCGGCTACAGCACCCC	230
Db	260	CAACAACTGTTGCTTACAGGCACAACGACAACAACACCTTAATTCACAACAACAACA	201
Qy	231	CTGGGGTATTTTGATTTTAAACAGATTCGACTGCGCACTTTTCCACCGTCACTGGGAGCG	290
Db	200	CTGTGTCTACAGGCACAACAACAACAACAACACTCTGTATTCAACAACAACAACAACATG	141
Qy	291	ACTCATCAACAACAATTTGGGGATTTCGGGCCAAGAGAGACTCAACTTCAAACTCTTCAACAT	350
Db	140	TTGCAGCAGGCAACAACAACGACAACAACAACCTGATTCACAACAACAATTTGCTACAG	81
Qy	351	CCAAGTCAAGGAGGTCAAGCAATGATGGGTCAACAACCATCGCTTAATTAACCTTACC	408
Db	80	GCAACAACAACAACAACAACAACAACTGTTCGAACAGGAACAACAAGCAACAACAACAC	23

RESULT 10	CO465395	801 bp	linear	EST 08-JUN-2005
LOCUS	MZCC520039C06.g	Maize Endosperm cDNA Library	Zea mays	CDNA, mRNA
DEFINITION	CO465395	sequence.		
ACCESSION	CO465395			
VERSION	CO465395.1	GI:67039140		
KEYWORDS	EST.			
SOURCE	Zea mays			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
REFERENCE	1 (bases 1 to 801)			
AUTHORS	Verza, N.C., Silva, T.R., Cord-Neto, G., Nogueira, F.T.S., Fisch, P.H., De Rosa, V.E. Jr., Martins, M.M., Vettore, A.L., da Silva, P.R. and Aruda, P.			
TITLE	Endosperm-preferred expression of maize genes as revealed by transcriptome-wide analysis of expressed sequence tags			
JOURNAL	Plant Mol. Biol.			
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br.			

FEATURES

/organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="F-352 near isogenic line"
 /db_xref="taxon:4577"
 /sex="hermaphrodite"
 /tissue_type="endosperm"
 /dev_stage="multiple stages (10 to 25 days after
 pollination, see publication for more information)"
 /lab_host="E. coli DH10B"
 /clone_lib="Maize Endosperm cDNA Library"
 Note: Organ: seed; Vector: pSPORT1; Site 1: SalI; Site 2:
 NotI; Plant Material and RNA Isolation: Field grown maize
 plants from inbred line F352 were used. Ears were
 harvested at 10, 15, 20 and 25 days after pollination
 (DAP), seeds were cut from the ear and the upper third of
 the endosperms, containing only endosperm, aleurone and
 pericarpal tissues, was removed, frozen in liquid nitrogen
 and stored at -80C. Frozen endosperms were pulverized in
 liquid nitrogen and total RNA was isolated according the
 method of Manning (9). Poly(A)+RNA was isolated using
 Oligotex-dT. cDNA libraries were constructed using
 Superscript Plasmid System for cDNA Synthesis and Plasmid
 Cloning Kit as described in Vettore, et al., (2001). The
 libraries that made SUGEST. Genet Mol Biol 24: 1-7. cDNAs
 ranging from 500 to 800 bp in size were assigned as short
 libraries (S10, S15, S20), and cDNAs >800 were assigned as
 long libraries (L10, L15, M15, N15, L20, L25). Unamplified
 libraries were plated and individual colonies picked and
 transferred to 96 well plates containing liquid Circie
 Grow (CG) medium supplemented with 100 mg/L of ampicillin
 and 8% glycerol. Three copies of each cDNA clone were
 stored at -80C. Additional information can be found in :
 Verza, N.C., Silva, T.R., Cord-Neto, G., Nogueira, F.T.S.,
 Fisch, P.H., De Rosa Jr., V.E., Martins, M.M., Vettore,
 A.L., da Silva, F.R. and Arruda, P. (2005)
 Endosperm-preferred expression of maize genes as revealed
 by transcriptome-wide analysis of expressed sequence tags.
 Plant Molecular Biology (in press//)

ORIGIN

Query Match	2.7%	Score 43.4;	DB 7;	Length 801;
Best Local Similarity	46.3%	Pred. No. 0.73;		
Matches 143;	Conservative	0;	Mismatches 166;	Indels 0;
Gaps 0;				

Qy	102	CGACAGAGTCTATCACCAACAGCACCGCACCTGGGGCTTTGCCACCTCTACAATAAACCACTT	161
Db	62	CAACAAAGCCATGAGCAGCAACATCATCCACAACAACATCATCCACAAAACAACAACAT	121
Qy	162	CTACAAGCAATCTCAGTGTCTTCAAGGGGGCCAGCAACGACAACCACTACTTCGGCTA	221
Db	122	CAACCCACCACCAACATCACCAGCAGCAACCAACAACAACAACATCACCAA	181
Qy	222	CAGCACCCCTGGGGTATTTGATTTCAACAGATTCACATGCCACTTTTTCACCACTGA	281
Db	182	CAATCAAGGGCCATGTGCAACAACAACGACAGAGCCATGAGCAACCAAGGACAGAGC	241
Qy	282	CTGCGAGGACTCATCAACAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACCT	341
Db	242	CATGAGCAACAATCAACAACAATTTCCAGGTCATGACAAGCAGCAACCAACAACAG	301
Qy	342	CTTCAACATCCAAGTCAAGGAGGTCAAGCGAATGATGGCGTCAACAACCATCCTAATAA	401
Db	302	CCTCAGCAATATCAGCAGGGCCAGGAAAATCACACAGCAACAATGTCTATGCCAGGAG	361
Qy	402	CCTTACCAG	410
Db	362	CAGCAACAG	370

RESULT 11

CO048199	748 bp	mRNA	linear	EST 10-JUN-2004
LOCUS				
DEFINITION	Lr PAHCF_66A08_SKplus Earthworm Fluorantene Exposure Library			
	Lumbricus rubellus cDNA clone Lr_PAHCF_66A08, mRNA sequence.			
ACCESSION	CO048199			
VERSION	CO048199.1			
KEYWORDS	GI:48588353			
SOURCE	EST.			
ORGANISM	Lumbricus rubellus (humus earthworm)			
	Lumbricus rubellus			
	Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;			
	Lumbricina; Lumbricidae; Lumbricus.			
REFERENCE	1 (bases 1 to 748)			
AUTHORS	Chaseley, J., Hedley, B.A., Morgan, J.C., Sturzenbaum, S., Kille, P. and Blaxter, M.			
TITLE	The lumbricus rubellus EST program - Sequences from a Fluorantene Exposure library			
JOURNAL	Unpublished (2003)			
COMMENT	Contact: Jennifer Chaseley			
	BICOSI 1			
	Cardiff University			
	Main College, Museum Avenue, Cardiff, CF11 3TL, UK			
	Tel: +44 2920876680			
	Fax: +44 2920874305			
	Email: chaseley@cardiff.ac.uk, Kille@cardiff.ac.uk			

FEATURES

```

1. 748
/organism="Lumbricus rubellus"
/mol_type="mrna"
/db_xref="taxon:35632"
/clone="dr PAHCF 66A08"
/tissue_type="Whole worm"
/dev_stage="Adult"
/clone_lib="Earthworm Fluorant
/notes="Vector: pBluescript II
prepared using protocols given
(Stratagene)."

```

ORIGIN

Query Match 2.7%; Score 43.2; DB 7; Length 748;
Best Local Similarity 51.6%; Pred. No. 0.82;

[illegible]

RESULT 12
BE556771
LOCUS
DEFINITION
BE556771 415 bp mRNA linear EST 30-AUG-2000
fr95c07.y1 Zebrafish Research Genomics C32 fin Danio rerio cdna 5'
similar to contains element TAR1 repetitive element ;, mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BE556771
BE556771.1
GI:9821261
EST.
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 415)
Clark, M., Johnson, S. L., Lebrach, H., Lee, R., Li, F., Marra, M.,
Edgy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Parson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
WashU Zebrafish EST Project 1998
Unpublished (1998)

TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: Ning Wu. cDNA Library Arrayed by:
Research Genetics. DNA Sequencing by: Washington University
Sequencing Center Clone distribution: Research Genetics web
address: <http://www.researchgenetics.com/>
Seq primer: T3 ET from Amersham
High quality sequence stop: 401.

```

FEATURES
    source
        Location/Qualifiers
            1..415
                /organism="Danio rerio"
                /mol_type="mRNA"
                /db_xref="taxon:7955"
                /tissue_type="pin"
                /lab_host="GeneHogs (HS996, a phage-resistant isolate of DH10B)"
                /clone_lib="Zebrafish Research Genetics C32 fin"
                /note="Vector: pVT73D-Pac with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was prepared from zebrafish(C32)-fin, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pVT73 vector. Library is non-normalized. Library was constructed by Ning Wu. NOTE: This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info.llnl.gov) for further information"

```

ORIGIN		Query Match	2.7%	Score 42.8;	DB 2;	Length 415;
		Best Local Similarity	54.4%;	Pred. No. 0.86;		
		Matches 86;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0;
Qy	111	CATCACACAGACCGGACCTTGGGGCTTGGCCACCTACAAATACACACCTCTACAAACA	170			
Db	41	CATCAGCAACTACCTTACAATGTCATCTTTCATCAACTACCTTACAACCTGGCGGCTTCAG	100			
Qy	171	AATCTCCAGTGTCTTCAACGGGGGGCCAGCAACGACCACTACTTGGGCTTACAGCACCCC	230			
Db	101	CAACTACACCTCAAACTGGGGCTTCAGCAACTACACCCCAAACTGGCGGCTTCAGCAACTA	160			
Qy	231	CTGGGGGTATTTTGATTTTCAACAGATTCCTACTGCCACT	268			
Db	161	CACCTTCAACTGGGGCTTCAGCAACTACACCTTCAACT	198			

RESULT 13					
CD920707					
LOCUS	CD920707	668 bp	mRNA	linear	EST 15-JUL-2003
DEFINITION	G808.118C06F010910 G608 Triticum aestivum cDNA clone G608118C06, mRNA sequence.				

CD920707.1	GI:32768471
CD920707.1	GI:32768471
EST.	
KEYWORDS	
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.
REFERENCE	1 (bases 1 to 668)

AUTHORS
Genopianté.
TITLE
Genopianté, a major partnership french program in plant genomics
JOURNAL
Unpublished (2003)
COMMENT
Contact: Genopianté
Genopianté
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genopianté' (<http://www.genopianté.com>
and <http://genopianté-info.infobiogen.fr>).

```

FEATURES
source
1. .668
/organism="Triticum aestivum"
/mol type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G60818C06"
/tissue type="grain (608 degrees per day after
pollination)"
/clone lib="G608"

```

ORIGIN	Query Match	2.7%; Score 42.8; DB 6; Length 668;
	Best Local Similarity 49.5%; Pred. No. 1;	
	Matches 110; Conservative 0; Mismatches 112; Indels 0; Gaps 0	
Qy	84 CGATTCCACATGGCTGGGCGCAGAGTCA	CACCACGACCGCACCTGGGGCTTGGC 143
Db	302 CAATCCCCGAGCAACCNCAATTC	CCCCNACCCNACCAACCAACACCCAC 361
Qy	144 CACCTACAAATACCACTCTTACAAGCA	AATCTCGAGTGGCTTCAAGGGGGCCAGCA 203
Db	362 CACCACCAATTAACAACACAAACAACA	CAACAACAACACACAAACAACACAA 421
Qy	204 CAACCACTACTTCGGCTACACGCCCT	CGCTGGGGTATTTTCATTTCAACAGATTC 263
Db	422 CAACCAAGCAACAACAACACGACNACA	CAGCAACAACAGCAACAGCAACAGCA 481
Qy	264 CCACTTTTCAACAGTGACTGGCAGCG	ACTCATCAACAACAA 305

```

Db      482 CAACAAATCCCCCAACAAACAATCCCGGAGCAACCAAA 523
|||||
BQ449242      435 bp      mRNA      linear      EST 29-MAY-2002
LOCUS      faa44912.y1 zebrafish fin day3 regeneration Danio rerio cDNA clone
DEFINITION      IMAGE:5912254 5' similar to contains element MER1 repetitive
                element /, mRNA sequence.
ACCESSION      BQ449242
VERSION      BQ449242.1 GI:21252354
KEYWORDS      EST.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
                Cypriniformes; Cyprinidae; Danio.
REFERENCE      1 (bases 1 to 435)
AUTHORS      Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,
                Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T.,
                Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
                Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
                Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
                Waterston,R. and Wilson,R.
                WashU Zebrafish EST Project 1998
TITLE      WashU Zebrafish EST Project 1998
JOURNAL      Unpublished (1998)
COMMENT      Contact: Stephen L. Johnson
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel: 314 286 1800
                Fax: 314 286 1810
                Email: zbrfish@wustl.edu
                cDNA Library Preparation: Raymond Lee. cDNA Library Arrayed by:
                Matthew Clark. DNA Sequencing by: Washington University Genome
                Sequencing Center Clone distribution: Genome Systems, St. Louis,
                Missouri (web address: www.genomesystems.com) (email contact:
                info@genomesystems.com) and Research Genetics, Huntsville, Alabama
                (web address: www.regen.com) (email contact: info@regen.com) and
                ResourceCentrumPrimarDatenbank, Berlin, Germany (web address:
                www.rzpd.de)
FEATURES      source
                1..435
                /organism="Danio rerio"
                /mol_type="mRNA"
                /db_xref="taxon:7955"
                /clone="IMAGE:5912254"
                /sex="mixed male and female"
                /tissue_type="3 day fin regenerates"
                /lab_host="E. coli XL0LR"
                /clone_lib="zebrafish fin day3 regeneration"
                /notes="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; 1st
                strand cDNA primed with (GA)10ACTAGTCTCGAG(T)18, followed
                by second strand synthesis, and ligated to 5' adapter
                (5')-aattcgacagag-3', 3'-gccgtgctc-5'. cDNA was cloned
                directionally (EcoRI/XhoI) into Stratagene Zap express
                lambda phage arms. Mass invivo excision done to obtain
                inserts in pBK-CMV phagemid."
ORIGIN
Query Match      2.6%; Score 42.2; DB 5; Length 435;
Best Local Similarity 53.3%; Pred. No. 1.3;
Matches 89; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY      102 CGACAGAGTCATCACCAAGCAGCCGCGCTTGGCGCTTCCCACTACAAATACCACT 161
Db      181 CAACCTGGCGCTTCAGCAACTACACCTACAACTCGCGCTTCAGCAACTACCACTG 240
QY      162 CTACAGCAAAATCTCCAGTGCTTCAACGGGGGCGCAGCAACGACCACTTTCGGCTA 221
Db      241 CGGCTTCAGCAACTACCTTCAACTGCGGCTTCAGCAACTACCTACCACTGCGAGCTT 300

RESULT 14
BQ449242      450 bp      mRNA      linear      EST 08-MAR-2002
LOCUS      BQ449242 Dictyostelium discoideum cDNA library, CF Dictyostelium
DEFINITION      discoideum cDNA clone ddc57m07 5', mRNA sequence.
ACCESSION      BQ449242
VERSION      BQ449242.1 GI:19280680
KEYWORDS      EST.
SOURCE      Dictyostelium discoideum
ORGANISM      Dictyostelium discoideum
                Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE      1 (bases 1 to 450)
AUTHORS      Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE      Full length cDNA of Dictyostelium discoideum at the culmination
                stage
JOURNAL      Unpublished (2002)
COMMENT      Contact: Tadasu Shin-i
                Center For Genetic Resource Information
                National Institute of Genetics
                1111 Yata, Mishima, Shizuoka 411-8540, Japan
                Tel: 81-559-81-6856
                Fax: 81-559-81-6855
                Email: tshini@genes.nig.ac.jp.
                Location/Qualifiers
                1..450
                /organism="Dictyostelium discoideum"
                /mol_type="mRNA"
                /strain="AX4"
                /db_xref="taxon:44689"
                /clone="ddc57m07"
                /sex="mat A"
                /dev_stage="Culmination stage"
                /clone_lib="Dictyostelium discoideum cDNA library, CF"
ORIGIN
Query Match      2.6%; Score 42.2; DB 3; Length 450;
Best Local Similarity 46.0%; Pred. No. 1.3;
Matches 143; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
QY      102 CGACAGAGTCATCACCAAGCAGCCGCGCTTGGCGCTTCCCACTACAAATACCACT 161
Db      116 CAACCAACAACATCAACAACATCAACAACATCAACAACATCAACAACATCAACAACAT 175
QY      162 CTACAGCAAAATCTCCAGTGCTTCAACGGGGGCGCAGCAACGACCACTTTCGGCTA 221
Db      176 CAACAACAACATCAACAACAGCATATACAACAACAACAGCAACCAACCAACAGCAA 235
QY      222 CAGCACCCCTGGGGGTATTTTGATTTCAACAGATTTCCACTGCGCACTTTTCACCACTGA 281
Db      236 CCACAACAACAACATCAACAACAACAACAACAACAACAACAACAACAACAACAACA 295
QY      282 CTGGCAGCGAGCTCATCAACAACAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACT 341
Db      296 CAACAACAACCAACCAACAGCAACAACAACAACCAACCAACCAACCAACCAACAG 355
QY      342 CTTCAACATCCAAAGTCAAGGAGGTCAACAGCATGATGGCGTCAACCACTCGCTAATAA 401
Db      356 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 415
QY      402 CCTTACCAGCA 412
Db      416 CAACAAGATCA 426

Search completed: November 29, 2005, 14:17:59
Job time : 5096.96 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:15:02 ; Search time 228.137 Seconds
(without alignments)
12497.813 Million cell updates/sec

Title: US-10-696-900-1_COPY_2829_4432
Perfect score: 1604
Sequence: 1 atggcttcaggcggtgggc.....tacctaccgtccctgtga 1604

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 88878028 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1:	/cgn2_6/prodata/1/ina/1 COMB.seq:*
2:	/cgn2_6/prodata/1/ina/5 COMB.seq:*
3:	/cgn2_6/prodata/1/ina/6A COMB.seq:*
4:	/cgn2_6/prodata/1/ina/6B COMB.seq:*
5:	/cgn2_6/prodata/1/ina/H COMB.seq:*
6:	/cgn2_6/prodata/1/ina/PTUS COMB.seq:*
7:	/cgn2_6/prodata/1/ina/PP COMB.seq:*
8:	/cgn2_6/prodata/1/ina/RE COMB.seq:*
9:	/cgn2_6/prodata/1/ina/baCkfilese1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1604	100.0	1605	3	US-09-807-802A-16
2	1604	100.0	1800	3	US-09-807-802A-14
3	1604	100.0	2211	3	US-09-807-802A-12
4	1604	100.0	4718	3	US-09-807-802A-1
5	1604	100.0	7447	3	US-10-216-870-11
6	1512.8	94.3	4683	3	US-09-807-802A-19
7	1026.4	64.0	4072	3	US-09-770-315-4
8	1026.4	64.0	4679	3	US-10-038-972A-12
9	1026.4	64.0	7557	3	US-09-770-315-3
10	1026.4	64.0	8698	3	US-09-770-315-2
11	1021.4	63.7	8179	3	US-09-438-268-5
12	1004.2	62.6	4680	2	US-08-254-358-1
13	1004.2	62.6	4680	2	US-08-475-391-1
14	1004.2	62.6	4680	2	US-08-709-609-1
15	1004.2	62.6	4680	6	PCT-US95-07178-1
16	999.6	62.3	4681	3	US-09-807-802A-18
17	987	61.5	4675	3	US-09-782-378A-1
18	987	61.5	4675	3	US-09-782-378A-2
19	902.8	56.3	8151	3	US-09-438-268-2
20	588.2	36.7	7214	3	US-09-438-268-1
21	588.2	36.7	7744	3	US-10-216-870-14
22	587.2	36.6	1617	3	US-09-532-594B-19
23	587.2	36.6	1800	3	US-09-532-594B-17
24	587.2	36.6	2208	3	US-09-532-594B-5

25	587.2	36.6	4767	3	US-09-532-594B-1	Sequence 1, Appli
26	537.4	33.5	2264	3	US-09-533-427-8	Sequence 8, Appli
27	537.4	33.5	2264	3	US-09-533-427-9	Sequence 9, Appli
28	537.4	33.5	2307	3	US-09-533-427-7	Sequence 7, Appli
29	537.4	33.5	4652	3	US-09-533-427-1	Sequence 1, Appli
30	46	2.9	5511	3	US-08-928-361B-2	Sequence 2, Appli
31	46	2.9	5511	3	US-09-588-995A-2	Sequence 2, Appli
32	46	2.9	7334	3	US-08-928-361B-1	Sequence 1, Appli
33	46	2.9	7334	3	US-09-588-995A-1	Sequence 1, Appli
34	42.8	2.7	5163	3	US-08-700-651-1	Sequence 4, Appli
35	42.8	2.7	5163	3	US-08-928-361B-4	Sequence 4, Appli
36	42.8	2.7	5163	3	US-09-588-995A-4	Sequence 4, Appli
37	42.8	2.7	5318	3	US-08-700-651-2	Sequence 2, Appli
38	42.8	2.7	5318	3	US-08-928-361B-3	Sequence 3, Appli
39	42.8	2.7	5318	3	US-09-588-995A-3	Sequence 3, Appli
40	41.4	2.6	1723	3	US-10-104-047-814	Sequence 814, App
C 41	41	2.6	34662	3	US-09-902-540-1261	Sequence 1261, Ap
C 42	40.4	2.5	601	3	US-09-949-016-165983	Sequence 165983,
43	40.4	2.5	61461	3	US-09-949-016-16419	Sequence 16419, A
44	39.4	2.5	2271	3	US-09-438-268-3	Sequence 3, Appli
45	39.2	2.4	700	3	US-10-187-253E-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-807-802A-16
; Sequence 16, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1602)
; OTHER INFORMATION:
US-09-807-802A-16

Query Match					100.0%;	Score 1604;	DB 3;	Length 1605;
Best Local Similarity					100.0%;	Pred. No. 0;		
Matches 1604;					Conservative	0;	Mismatches	0; Gaps 0;
QY	1	ATGGCTTCAGGCGGTGGCGC	CAATGGCAGCAATAACGAAGCGCGCGAGTGGGT	60				
Db	1	ATGGCTTCAGGCGGTGGCGC	CAATGGCAGCAATAACGAAGCGCGCGAGTGGGT	60				
QY	61	AATGCTTCAGGAAATGGCAT	TTCGATTCACATGCTGGGCGACAGAGTCATCACCACC	120				
Db	61	AATGCTTCAGGAAATGGCAT	TTCGATTCACATGCTGGGCGACAGAGTCATCACCACC	120				
QY	121	AGCACCCGACCTGGGCGCTT	GCCACCTCAATACCACTCTTACAGCAAAATCTCCAGT	180				
Db	121	AGCACCCGACCTGGGCGCTT	GCCACCTCAATACCACTCTTACAGCAAAATCTCCAGT	180				
QY	181	GCTTCAACGGGGCCAGCAAC	AGCACTACTTCTGGGTACAGACCCCTCGGGGTAT	240				
Db	181	GCTTCAACGGGGCCAGCAAC	AGCACTACTTCTGGGTACAGACCCCTCGGGGTAT	240				

```
QY 241 TTTGATTTCAACAGATTCACACTGCCACTTTTTCACACGCTGACTGGCGGACTCATCAAC 300
Db 241 TTTGATTTCAACAGATTCACACTGCCACTTTTTCACACGCTGACTGGCGGACTCATCAAC 300
QY 301 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTTCAAACCTCTTCAAACATCAAGTCAAG 360
Db 301 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTTCAAACCTCTTCAAACATCAAGTCAAG 360
QY 361 GAGGTCAAGCAAGATGATGGCTCAACACCATCGCTTAATCACTTACCAAGCAAGTTCAA 420
Db 361 GAGGTCAAGCAAGATGATGGCTCAACACCATCGCTTAATCACTTACCAAGCAAGTTCAA 420
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGCGCACGAGGCTGC 480
Db 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGCGCACGAGGCTGC 480
QY 481 CTCCCTCCGTTCCCGCGGACGCTGTTTCATGATTCGCGCAATACGGCTACCTGACGCTCAAC 540
Db 481 CTCCCTCCGTTCCCGCGGACGCTGTTTCATGATTCGCGCAATACGGCTACCTGACGCTCAAC 540
QY 541 AATGGCAGCAAGCCGTGGGACGTTTCATCTTTTACTGCTCGAATATTTTCCCTTCTCAG 600
Db 541 AATGGCAGCAAGCCGTGGGACGTTTCATCTTTTACTGCTCGAATATTTTCCCTTCTCAG 600
QY 601 ATGCTGAGAACCGGCNAACAACTTACCTTCAGTACACTTTTGAGGAAGTGCCTTCCAC 660
Db 601 ATGCTGAGAACCGGCNAACAACTTACCTTCAGTACACTTTTGAGGAAGTGCCTTCCAC 660
QY 661 AGCAGCTACGCGCACAGCAGCAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
Db 661 AGCAGCTACGCGCACAGCAGCAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
QY 721 CTGTATTACTGACAGAACTCAAATCAGTCCGGAAGTGCCCAAAACAAAGCACTTGCTG 780
Db 721 CTGTATTACTGACAGAACTCAAATCAGTCCGGAAGTGCCCAAAACAAAGCACTTGCTG 780
QY 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACTCGGACCC 840
Db 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACTCGGACCC 840
QY 841 TGTATTCCGCGAGCGCGCTTTCTAAACAAAAACAGACAAACAAACAGCAATTTTACC 900
Db 841 TGTATTCCGCGAGCGCGCTTTCTAAACAAAAACAGACAAACAAACAGCAATTTTACC 900
QY 901 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
Db 901 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
QY 961 GCTATGGCTCTACAAAGACGACGACGACGATGCTTTTCCCATGAGCGGTGCATGATT 1020
Db 961 GCTATGGCTCTACAAAGACGACGACGACGATGCTTTTCCCATGAGCGGTGCATGATT 1020
QY 1021 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGTCATTTGGACAAATGTCATGATTACAGAC 1080
Db 1021 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGTCATTTGGACAAATGTCATGATTACAGAC 1080
QY 1081 GAAGAGGAAATTAAGGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 1081 GAAGAGGAAATTAAGGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
QY 1141 AATTTCCAGAGCAGCAGCAGACCTTGGACCGGAGATGTGCTATGGGAGCAATTA 1200
Db 1141 AATTTCCAGAGCAGCAGCAGACCTTGGACCGGAGATGTGCTATGGGAGCAATTA 1200
QY 1201 CTTGGCATGGTGGCAAGATAGAGACGTGTACTTCGAGGGTCCCATTTTGGGCCAAAAT 1260
Db 1201 CTTGGCATGGTGGCAAGATAGAGACGTGTACTTCGAGGGTCCCATTTTGGGCCAAAAT 1260
QY 1261 CTTCAACACAGATGGAACATTTTCAACCGTCTCTCTTTATGGGGCGCTTTTGGACTCAAGAAC 1320
Db 1261 CTTCAACACAGATGGAACATTTTCAACCGTCTCTCTTTATGGGGCGCTTTTGGACTCAAGAAC 1320
```

```
QY 1321 CGGCTCTCAGATCTCATCAAAAAACAGCGCTGTCTCTCGAATCTCTCCGCGGAGTTT 1380
Db 1321 CGGCTCTCAGATCTCATCAAAAAACAGCGCTGTCTCTCGAATCTCTCCGCGGAGTTT 1380
QY 1381 TCAGCTACAAAAGTTTGTCTTCAATTCATCAACCCCAATATCTCCACAGGACAAAGTGTGGAA 1440
Db 1381 TCAGCTACAAAAGTTTGTCTTCAATTCATCAACCCCAATATCTCCACAGGACAAAGTGTGGAA 1440
QY 1441 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 1500
Db 1441 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 1500
QY 1501 TCCAAATATGCAAAATCTGCCAAGCTGTTTACTGTGGACAAACATGGACTTTTATCT 1560
Db 1501 TCCAAATATGCAAAATCTGCCAAGCTGTTTACTGTGGACAAACATGGACTTTTATCT 1560
QY 1561 GAGCCTCCGCCCATTTGGCACCCGTTTACCTTACCCGTTCCCTGTGA 1604
Db 1561 GAGCCTCCGCCCATTTGGCACCCGTTTACCTTACCCGTTCCCTGTGA 1604
```

RESULT 2

```
US-09-807-802A-14
; Sequence 14, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1797)
; OTHER INFORMATION:
US-09-807-802A-14
```

Query Match 100.0%; Score 1604; DB 3; Length 1800;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCTTCAGGCGTGGCGCACCAATGCGAGCAATAACGAAGGCGCGACGAGTGGGT 60
Db 196 ATGGCTTCAGGCGTGGCGCACCAATGCGAGCAATAACGAAGGCGCGACGAGTGGGT 255
QY 61 AATGCTCAGGAAATTTGGCATTTGCCATTTCCATGGCTGGGCGACAGAGTCAATCACCAC 120
Db 256 AATGCTCAGGAAATTTGGCATTTGCCATTTCCATGGCTGGGCGACAGAGTCAATCACCAC 315
QY 121 AGCACCGCACCTGGCGCTTGGCCACCTTACATTAACCACTACTTTCGGCTACAGCACCCCTCGGGGTAT 180
Db 316 AGCACCGCACCTGGCGCTTGGCCACCTTACATTAACCACTACTTTCGGCTACAGCACCCCTCGGGGTAT 375
QY 181 GCTTCAACGGGGCGCAGCAACGACAAACCACTACTTTCGGCTACAGCACCCCTCGGGGTAT 240
Db 376 GCTTCAACGGGGCGCAGCAACGACAAACCACTACTTTCGGCTACAGCACCCCTCGGGGTAT 435
QY 241 TTTGATTTCAACAGATTTCACTGGCACTTTTCCACAGTGACTGGCAGGCACTCATCAAC 300
Db 436 TTTGATTTCAACAGATTTCACTGGCACTTTTCCACAGTGACTGGCAGGCACTCATCAAC 495
```

QY 301 AACAAATGGGGATTCCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360
DB 496 AACAAATGGGGATTCCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 355
QY 361 GAGGTACGACGAATGATGGCGTCAACAACATCGCTTAATACCTTACCGACGCGTTCAA 420
DB 556 GAGGTACGACGAATGATGGCGTCAACAACATCGCTTAATACCTTACCGACGCGTTCAA 615
QY 421 GTCTTCTCGGACTCGGAGTACCAAGCTTCCGTACGTCTCGGCTCTCGGCACCAAGGCTGC 480
DB 616 GTCTTCTCGGACTCGGAGTACCAAGCTTCCGTACGTCTCGGCTCTCGGCACCAAGGCTGC 675
QY 481 CTCCTCTCGGTTCCCGCGGACGTTGTTTCATGATTCGGAATACGCGTACCTGACGCTCAAC 540
DB 676 CTCCTCTCGGTTCCCGCGGACGTTGTTTCATGATTCGGAATACGCGTACCTGACGCTCAAC 735
QY 541 AATGGCAGCCAGCCGCTGGAGCGTTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 600
DB 736 AATGGCAGCCAGCCGCTGGAGCGTTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 795
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660
DB 796 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 855
QY 661 AGCAGCTACGCGCACAGCCGAGCCCTGGACCGCTGATGAATCTCTCATCGACCAATAC 720
DB 856 AGCAGCTACGCGCACAGCCGAGCCCTGGACCGCTGATGAATCTCTCATCGACCAATAC 915
QY 721 CTGTATTACTGAAACGAACCTAAATTCAGTCCGGAAGTGCCTTCAACAAAGGACTTGGTG 780
DB 916 CTGTATTACTGAAACGAACCTAAATTCAGTCCGGAAGTGCCTTCAACAAAGGACTTGGTG 975
QY 781 TTTAGCCGTGGGTCTCCAGCTGCGATGCTGTTTCAGCCCAAAACCTGGCTACCTGGACCC 840
DB 976 TTTAGCCGTGGGTCTCCAGCTGCGATGCTGTTTCAGCCCAAAACCTGGCTACCTGGACCC 1035
QY 841 TGTATTCCGCGAGCGCGCTTCTTAAACAAACACAGACAAACACAGCAATTTTACC 900
DB 1036 TGTATTCCGCGAGCGCGCTTCTTAAACAAACACAGACAAACACAGCAATTTTACC 1095
QY 901 TGGACTGGTGCTTCAAAATTAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960
DB 1096 TGGACTGGTGCTTCAAAATTAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 1155
QY 961 GCTATGSCCTCACAAAAGACGACGAGCAAGTCTTCCCATGAGCGGTGTCATGATT 1020
DB 1156 GCTATGSCCTCACAAAAGACGACGAGCAAGTCTTCCCATGAGCGGTGTCATGATT 1215
QY 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAATGTTCATGATTACAGAC 1080
DB 1216 TTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAATGTTCATGATTACAGAC 1275
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTGGGACCGTGGCAGTC 1140
DB 1276 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTGGGACCGTGGCAGTC 1335
QY 1141 AATTTCCAGACGACGACACAGACCCCTGCCACCGGAGATGTCATGTCATGGGACATTA 1200
DB 1336 AATTTCCAGACGACGACACAGACCCCTGCCACCGGAGATGTCATGTCATGGGACATTA 1395
QY 1201 CCTGGCATGGTGGCAAGATAGAGACGTGTACTCTGACGGGTCCCATTTGGGCCAAAATT 1260
DB 1396 CCTGGCATGGTGGCAAGATAGAGACGTGTACTCTGACGGGTCCCATTTGGGCCAAAATT 1455
QY 1261 CCTCACACAGATGGACATTTTCAACCGCTCTCTCTTATAGGCGGCTTTGGAATCAAGAAC 1320
DB 1456 CCTCACACAGATGGACATTTTCAACCGCTCTCTCTTATAGGCGGCTTTGGAATCAAGAAC 1515
QY 1321 CCGCTCTCAGATCTCATCAAAAACAGCGCTGTTCTCGGATCTCTCGGCGGAGTTT 1380
DB 1516 CCGCTCTCAGATCTCTCATCAAAAACAGCGCTGTTCTCGGATCTCTCGGCGGAGTTT 1575
QY 1381 TCAGCTACAAAGTTTGTCTTCAATTCATCAACCAATTAATCTCCACAGGACAAAGTGTGGAA 1440

DB 1576 TCAGCTACAAAGTTTGTCTTCAATTCATCAACCAATATCTCACAGGACAAGTGTGGAA 1635
QY 1441 ATTGAATGGAGCTGCAGAAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
DB 1636 ATTGAATGGAGCTGCAGAAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1695
QY 1501 TCAAAATTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTTTTACT 1560
DB 1696 TCAAAATTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTTTTACT 1755
QY 1561 GAGCCTCGCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGTA 1604
DB 1756 GAGCCTCGCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGTA 1799

RESULT 3

US-09-807-802A-12
; Sequence 12, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilison, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 2211
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2208)
; OTHER INFORMATION:
US-09-807-802A-12

Query Match 100.0%; Score 1604; DB 3; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGGCGTGGCGCACCAATGGCAGACAATAAGAGGCGCGCAGGAGTGGGT 60
DB 607 ATGGCTTCAGGGCGTGGCGCACCAATGGCAGACAATAAGAGGCGCGCAGGAGTGGGT 666
QY 61 AATGCCTCAGGAAATTTGGCATTTCCGATTTCCACATGGCTGGGCGACAGAGTCATCACACC 120
DB 667 AATGCCTCAGGAAATTTGGCATTTCCGATTTCCACATGGCTGGGCGACAGAGTCATCACACC 726
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTCAATAACACCTCTACAAGCAAAATCTCCAGT 180
DB 727 AGCACCCGACCTGGGCGCTTGGCCACCTCAATAACACCTCTACAAGCAAAATCTCCAGT 786
QY 181 GCTTCAACGGGGCCAGCAACACCACTACTCTCGGCTACAGCACCCCTCGGGGTAT 240
DB 787 GCTTCAACGGGGCCAGCAACACCACTACTCTCGGCTACAGCACCCCTCGGGGTAT 846
QY 241 TTTGATTTCAACAGATTCACCTGCGCACTTTTACACAGTGACTGGGAGGAGTCAATCAAC 300
DB 847 TTTGATTTCAACAGATTCACCTGCGCACTTTTACACAGTGACTGGGAGGAGTCAATCAAC 906
QY 301 AACAAATGGGGAATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360
DB 907 AACAAATGGGGAATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 966
QY 361 GAGGTACAGCAAGTAATGATGGCGTGCACAAACCATCGCTAATAACCTTACAGCACGGTTCAA 420

Db 967 GAGGTACGACGAAATGATGGCGTCACACCATCGCTAATACCTTACCAGACGGTTCAA 1026
QY 421 GTCTTTCGGACTCGGAGTACACAGCTTCGGTACGTCCTCGGCTTCGGGACCGAGGCTGC 480
Db 1027 GTCTTTCGGACTCGGAGTACACAGCTTCGGTACGTCCTCGGCTTCGGGACCGAGGCTGC 1086
QY 481 CTCCTCTCGTTCCCGCGGACGCTGTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 540
Db 1087 CTCCTCTCGTTCCCGCGGACGCTGTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 1146
QY 541 AATGCGACGAAGCGTGGGAGCTTTCATCTTTTACTGCTCGGAATATTTCCCTTCTCAG 600
Db 1147 AATGCGACGAAGCGTGGGAGCTTTCATCTTTTACTGCTCGGAATATTTCCCTTCTCAG 1206
QY 601 ATGCTGAGAACGGGCAACATTTTACCTTCAGCTACACTTTTGAGGAAGTGCCTTTCCAC 660
Db 1207 ATGCTGAGAACGGGCAACATTTTACCTTCAGCTACACTTTTGAGGAAGTGCCTTTCCAC 1266
QY 661 AGCAGCTACGCGCAGCAGCAGAGCTGACCGGCTGATGATCCTCTCATCGACCAATAC 720
Db 1267 AGCAGCTACGCGCAGCAGCAGAGCTGACCGGCTGATGATCCTCTCATCGACCAATAC 1326
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGGACTTGTG 780
Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGGACTTGTG 1386
QY 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCGACCC 840
Db 1387 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCGACCC 1446
QY 841 TGTATTCCGCGACAGCGGCTTTCTAAACAAAAACAGACAAACAGAGCAATTTTACC 900
Db 1447 TGTATTCCGCGACAGCGGCTTTCTAAACAAAAACAGACAAACAGAGCAATTTTACC 1506
QY 901 TGGACTGGTGTCTTCAAAATATTAACCTCAATGGCGTGAATCCATCATCAACCTCGCACT 960
Db 1507 TGGACTGGTGTCTTCAAAATATTAACCTCAATGGCGTGAATCCATCATCAACCTCGCACT 1566
QY 961 GCTATGGCTCTCACAAAGACGACGAAGAAGTCTTTTCCCATGAGCGGTGTATGATT 1020
Db 1567 GCTATGGCTCTCACAAAGACGACGAAGAAGTCTTTTCCCATGAGCGGTGTATGATT 1626
QY 1021 TTTGGAAAGAGCGCGGAGCTTCAACATCTGCAATTTGGCAATGTCTATGATTACAGAC 1080
Db 1627 TTTGGAAAGAGCGCGGAGCTTCAACATCTGCAATTTGGCAATGTCTATGATTACAGAC 1686
QY 1081 GAAGAGGAATTTAAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 1687 GAAGAGGAATTTAAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746
QY 1141 AATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCTATGAGGAGATT 1200
Db 1747 AATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCTATGAGGAGATT 1806
QY 1201 CTGCGATGGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAATT 1260
Db 1807 CTGCGATGGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAATT 1866
QY 1261 CTTCAACAGATGGACACTTTTACCCGCTCTCTCTTTATGGCGGCTTTGAGCTCAAGAAC 1320
Db 1867 CTTCAACAGATGGACACTTTTACCCGCTCTCTCTTTATGGCGGCTTTGAGCTCAAGAAC 1926
QY 1321 CGCCCTCTCAGATCTCATCAAAAACAGCCCTGTTCCTGCGAATCCTCCGCGGAGTTT 1380
Db 1927 CGCCCTCTCAGATCTCATCAAAAACAGCCCTGTTCCTGCGAATCCTCCGCGGAGTTT 1986
QY 1381 TCAGCTACAAGTTTCTTATTTCATCAACCAATCTCCAGGACAAAGTGTGGAA 1440
Db 1987 TCAGCTACAAGTTTCTTATTTCATCAACCAATCTCCAGGACAAAGTGTGGAA 2046
QY 1441 ATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGACGATACACA 1500

Db 2047 ATTGAATGGGAGCTGCAGAAAGAAACACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 2106
QY 1501 TCCAATTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGACAAACAATGGACTTTTACT 1560
Db 2107 TCCAATTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGACAAACAATGGACTTTTACT 2166
QY 1561 GAGCTCGCCCCATTGGCACCCTGTTACTTACCCTGCCCTGTA 1604
Db 2167 GAGCTCGCCCCATTGGCACCCTGTTACTTACCCTGCCCTGTA 2210

RESULT 4
US-09-807-802A-1
; Sequence 1, Application US/09807802A
; Patent No.: 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN-031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:
US-09-807-802A-1

Query Match 100.0%; Score 1604; DB 3; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGCGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 60
Db 2829 ATGGCTTCAGGCGGTGGCGCACCAATGCGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 2888
QY 61 AATGCTCAGGAATTTGGCATTTGCCATTTGCCATTTCCACATGGCTGGGCGACAGAGTCAATCACC 120
Db 2889 AATGCTCAGGAATTTGGCATTTGCCATTTGCCATTTCCACATGGCTGGGCGACAGAGTCAATCACC 2948
QY 121 AGCACCCGACCTGGCGCTTGGCCACCTACAAATACCACTTACCAAGCAATCTCCAGT 180
Db 2949 AGCACCCGACCTGGCGCTTGGCCACCTACAAATACCACTTACCAAGCAATCTCCAGT 3008
QY 181 GCTTCAAACGGGGCGCAGCAACGACCAACCACTTCTGGCTACAGCACCCCTTGGGGGTAT 240
Db 3009 GCTTCAAACGGGGCGCAGCAACGACCAACCACTTCTGGCTACAGCACCCCTTGGGGGTAT 3068
QY 241 TTTGATTTCAACAGATTCACATGCGCATTTTTCACATGGCTGGGCGAGGAGTCAATCAAC 300
Db 3069 TTTGATTTCAACAGATTCACATGCGCATTTTTCACATGGCTGGGCGAGGAGTCAATCAAC 3128
QY 301 AACAAATGGGATTCGGGCCCAAGAGACTCAACTTTCAACTTTCAACTTTCAACTTTCAACTTTCAAGTCAAG 360
Db 3129 AACAAATGGGATTCGGGCCCAAGAGACTCAACTTTCAACTTTCAACTTTCAACTTTCAAGTCAAG 3188
QY 361 GAGGTACAGCAAGAAATGATGGGTCAACCAACCATCGCTAACTTAACTTACCAGCAACGGTTCAA 420

Db 3189 GAGGTCAACGAGTATGGCGTCAACAACATCGCTAAATAAACCCTTACCGACGCGTTCAA 3248
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCCTCGGCTCTCGCAGCACCAGGCGTGC 480
Db 3249 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCCTCGGCTCTCGCAGCACCAGGCGTGC 3308
QY 481 CTCCCTCCGTTCCCGCGGACGCGTTCATGATTCGCAATAACGCTACCTGACGCTCAAC 540
Db 3309 CTCCCTCCGTTCCCGCGGACGCGTTCATGATTCGCAATAACGCTACCTGACGCTCAAC 3368
QY 541 AATGGCAGCAGCAGCGTGGAGCGTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 600
Db 3369 AATGGCAGCAGCAGCGTGGAGCGTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 3428
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660
Db 3429 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 3488
QY 661 AGCAGCTACCGGCACAGCAGCGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 3489 AGCAGCTACCGGCACAGCAGCGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 3548
QY 721 CTGTATTACTGAGACGAACTCAAAATCAGTCCGGAAGTGCCTCAACAGGACTTGCCTG 780
Db 3549 CTGTATTACTGAGACGAACTCAAAATCAGTCCGGAAGTGCCTCAACAGGACTTGCCTG 3608
QY 781 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGGTACCTGGACCC 840
Db 3609 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGGTACCTGGACCC 3668
QY 841 TGTATTCCGAGCAGCGCGCTTTCTAAACAAAAACAGACAAACAAACGCAATTTTACC 900
Db 3669 TGTATTCCGAGCAGCGCGCTTTCTAAACAAAAACAGACAAACAAACGCAATTTTACC 3728
QY 901 TGGACTGGTGTCTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960
Db 3729 TGGACTGGTGTCTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 3788
QY 961 GCTATGGCTCTACAAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTATGATT 1020
Db 3789 GCTATGGCTCTACAAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTATGATT 3848
QY 1021 TTTGAAAAAGAGAGCGCCGAGCTTCAAAACACTGCAATTTGGACAATGTATGATTACAGAC 1080
Db 3849 TTTGAAAAAGAGAGCGCCGAGCTTCAAAACACTGCAATTTGGACAATGTATGATTACAGAC 3908
QY 1081 GAAGAGAAATTAAGGCACTAACCTGTGTGGCCACCGAAAGATTTGGGACCGTGGGAGTC 1140
Db 3909 GAAGAGAAATTAAGGCACTAACCTGTGTGGCCACCGAAAGATTTGGGACCGTGGGAGTC 3968
QY 1141 AATTTCCAGAGCAGCAGCAGACCTTGGACCGGAGATGTGCATGCTATGGGAGCATTA 1200
Db 3969 AATTTCCAGAGCAGCAGCAGACCTTGGACCGGAGATGTGCATGCTATGGGAGCATTA 4028
QY 1201 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAAAT 1260
Db 4029 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAAAT 4088
QY 1261 CCTCACACAGATGGACACTTTTACACCGTCTCTTATTTGGGCGGCTTTGGACTCAAGAAC 1320
Db 4089 CCTCACACAGATGGACACTTTTACACCGTCTCTTATTTGGGCGGCTTTGGACTCAAGAAC 4148
QY 1321 CCGCTCTCCAGATCTCATCAAAAACACGCTGTCTTCGCGAATCTCTCGGCGGAGTTT 1380
Db 4149 CCGCTCTCCAGATCTCATCAAAAACACGCTGTCTTCGCGAATCTCTCGGCGGAGTTT 4208
QY 1381 TCAGCTACAAGTTTGTCTTCAITCATCAACCCAAATCTCCAAGGACAAGTGAAGTGGAA 1440
Db 4209 TCAGCTACAAGTTTGTCTTCAITCATCAACCCAAATCTCCAAGGACAAGTGAAGTGGAA 4268
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAAACAGACGCTGGAAATCCCGAAGTGCAGTACACA 1500
Db 4269 ATTGAATGGAGCTGCAGAAAGAAAAACAGACGCTGGAAATCCCGAAGTGCAGTACACA 4328

QY 1501 TCCAATTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTTTATACT 1560
Db 4329 TCCAATTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTTTATACT 4388
QY 1561 GAGCCTCGCCCATTTGGCAGCCGCTTACCTTACCCTGCCCTCTGA 1604
Db 4389 GAGCCTCGCCCATTTGGCAGCCGCTTACCTTACCCTGCCCTCTGA 4432

RESULT 5
US-10-216-870-11
; Sequence 11, Application US/10216870
; Patent No. 6723551
; GENERAL INFORMATION:
; APPLICANT: KOTIN, ROBERT M
; APPLICANT: URABE, MASASHI
; APPLICANT: DING, CHUAN-TIAN
; TITLE OF INVENTION: PRODUCTION OF ADENO-ASSOCIATED VIRUS IN INSECT CELLS
; FILE REFERENCE: 402133
; CURRENT APPLICATION NUMBER: US/10/216,870
; CURRENT FILING DATE: 2002-08-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 7447
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-216-870-11

Query Match 100.0%; Score 1604; DB 3; Length 7447;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGCGCGTGGCGCAACAATGGCAGACAATAACGAAGGCGCCACGAGTGGGT 60
Db 5227 ATGGCTTCAGCGCGTGGCGCAACAATGGCAGACAATAACGAAGGCGCCACGAGTGGGT 5286
QY 61 AATGCTCAGGAAATTTGGCATTCACATGCTGGCGAGACAGTTCATCACACC 120
Db 5287 AATGCTCAGGAAATTTGGCATTCACATGCTGGCGAGACAGTTCATCACACC 5346
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACATAAACAACCTCTCAAGCAAAATCTCCAGT 180
Db 5347 AGCACCCGACCTGGGCGCTTGGCCACCTACATAAACAACCTCTCAAGCAAAATCTCCAGT 5406
QY 181 GCTTCAACGGGGCCAGCAACCACTACTTTCGGCTACAGCAACCCCTCGGGGGGTAT 240
Db 5407 GCTTCAACGGGGCCAGCAACCACTACTTTCGGCTACAGCAACCCCTCGGGGGGTAT 5466
QY 241 TTTGATTTCAACAGATTCACCTGCGCACTTTTCCACCGTGACTGGCAGCGACTCATCAAC 300
Db 5467 TTTGATTTCAACAGATTCACCTGCGCACTTTTCCACCGTGACTGGCAGCGACTCATCAAC 5526
QY 301 AACAAATTTGGGATTCGGGCGCCAAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360
Db 5527 AACAAATTTGGGATTCGGGCGCCAAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 5586
QY 361 GAGGTCAACGAATGATGGGTCAACCATCGCTTAATAAACCCTTACAGCAGCGTTCAA 420
Db 5587 GAGGTCAACGAATGATGGGTCAACCATCGCTTAATAAACCCTTACAGCAGCGTTCAA 5646
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCCGTAAGTTCGCTCGGCTCTCGGACACGAGGCTGC 480
Db 5647 GTCTTCTCGGACTCGGAGTACAGCTTCCGTAAGTTCGCTCGGCTCTCGGACACGAGGCTGC 5706
QY 481 CTCCCTCCGTTCCCGGCGAGCGTGTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 540
Db 5707 CTCCCTCCGTTCCCGGCGAGCGTGTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 5766
QY 541 AATGGCAGCAGCAGCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTTCCCTTCTCAG 600

```
Db      5767 AATGCGCAGCGCGTGGAGGTTTCATCTTTTACTGCTGGAAATATTTCCCTTCTAG 5826
QY      601 ATGCTGAGAACGGGCAACATTTTACCTTTCAGTACACCTTTTGAGGAAGTGCCTTTCCAC 660
Db      5827 ATGCTGAGAACGGGCAACATTTTACCTTTCAGTACACCTTTTGAGGAAGTGCCTTTCCAC 5886
QY      661 AGCAGTACGCCACAGCAGCAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db      5887 AGCAGTACGCCACAGCAGCAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 5946
QY      721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGACTTGCTG 780
Db      5947 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGACTTGCTG 6006
QY      781 TTTAGCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACTCGGACCC 840
Db      6007 TTTAGCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACTCGGACCC 6066
QY      841 TGTTATCGGCAGAGCGGCTTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGCACT 960
Db      6067 TGTTATCGGCAGAGCGGCTTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGCACT 6126
QY      901 TGAGCTGGTGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGCACT 960
Db      6127 TGAGCTGGTGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGCACT 6186
QY      961 GCTATGGCTCTCACAAAGACGACGAAGTCTTCCATGAGCGGTGTGATGTT 1020
Db      6187 GCTATGGCTCTCACAAAGACGACGAAGTCTTCCATGAGCGGTGTGATGTT 6246
QY      1021 TTTGNAAGAGAGCGCGGAGCTTCAACACATGCTATGGCAATGTCTATGATTAAGAC 1080
Db      6247 TTTGNAAGAGAGCGCGGAGCTTCAACACATGCTATGGCAATGTCTATGATTAAGAC 6306
QY      1081 GAAGAGGAATTTAAAGCCTTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db      6307 GAAGAGGAATTTAAAGCCTTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 6366
QY      1141 AATTTCCAGAGCAGCAGCAGACACCTGCGACCGGAGATGTGCTATGGGAGCATTA 1200
Db      6367 AATTTCCAGAGCAGCAGCAGACACCTGCGACCGGAGATGTGCTATGGGAGCATTA 6426
QY      1201 CTTGGCATGGTGGCAAGATAGAGACGTGTACTGCGAGGTTCCCATTTTGGGCCAAAATT 1260
Db      6427 CTTGGCATGGTGGCAAGATAGAGACGTGTACTGCGAGGTTCCCATTTTGGGCCAAAATT 6486
QY      1261 CCTCACAGATGGACATTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db      6487 CCTCACAGATGGACATTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 6546
QY      1321 CGCCTCTCTCAGATCCTCATCAAAAACACGCTGTGTTCTTGGCAATCCTCCGCGGAGTTT 1380
Db      6547 CGCCTCTCTCAGATCCTCATCAAAAACACGCTGTGTTCTTGGCAATCCTCCGCGGAGTTT 6606
QY      1381 TCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCACAGGACAAGTGTGGAA 1440
Db      6607 TCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCACAGGACAAGTGTGGAA 6666
QY      1441 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA 1500
Db      6667 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA 6726
QY      1501 TCAATATTGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTATCT 1560
Db      6727 TCAATATTGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTATCT 6786
QY      1561 GAGCCTCGCCCAATTTGGGACCGGTTACCTTACCGTCCCTGTGA 1604
Db      6787 GAGCCTCGCCCAATTTGGGACCGGTTACCTTACCGTCCCTGTGA 6830
```

```
US-09-807-802A-19
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVFN.03105A
; CURRENT APPLICATION NUMBER: US/09/807.802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19
```

```
Query Match      94.3%; Score 1512.8; DB 3; Length 4683;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY      1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGAGTGGGT 60
Db      2814 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGAGTGGGT 2873
QY      61 AATGCTTCAGGAATTTGGCATTTGGATTCCACATGGCTTGGGGGACAGAGTATCATCACACC 120
Db      2874 AATGCTTCAGGAATTTGGCATTTGGATTCCACATGGCTTGGGGGACAGAGTATCATCACACC 2933
QY      121 AGCACCCGACCTGGGCTTGGCCACCTACAAATTAACCACTCTCAACGCAAAATCTCCAGT 180
Db      2934 AGCACCCGAAACATGGGCTTGGCCACCTTAAACCAACCTCTCAACGCAAAATCTCCAGT 2993
QY      181 GCTTCAACGGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 240
Db      2994 GCTTCAACGGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 3053
QY      241 TTTGATTTCAACAGATTCCACTGGCACCTTTTCAACACGTGACTGGCAGGACTCATCAAC 300
Db      3054 TTTGATTTCAACAGATTCCACTGGCACCTTTTCAACACGTGACTGGCAGGACTCATCAAC 3113
QY      301 AACAAATGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAACATCCAAGTCAAG 360
Db      3114 AACAAATGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAACATCCAAGTCAAG 3173
QY      361 GAGGTACAGACGAATGATGGGTCAACACCATCGCTAATAACCTTTACAGCACGGTTCAA 420
Db      3174 GAGGTACAGACGAATGATGGGTCAACACCATCGCTAATAACCTTTACAGCACGGTTCAA 3233
QY      421 GTCCTTCGGACTCGGAGTACAGCTTCGGTACGTCTCTCGGCTCTGGCGCACCGGCTGC 480
Db      3234 GTCCTTCGGACTCGGAGTACAGCTTCGGTACGTCTCTCGGCTCTGGCGCACCGGCTGC 3293
QY      481 CTCCTCTCGTTCGGCGGAGAGTTCATGTTTCCGCAATACGGCTACTGACGCTCAAC 540
Db      3294 CTCCTCTCGTTCGGCGGAGAGTTCATGTTTCCGCAATACGGCTACTGACGCTCAAC 3353
QY      541 AATGCGACCCAGCGGTGGGACGTTTCCTTTTACTGCTGGAATATTTCCCTTTCTCAG 600
Db      3354 AATGCGACCCAGCGGTGGGACGTTTCCTTTTACTGCTGGAATATTTCCCATCGCAG 3413
QY      601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACTTTTGAGGAAGTGCCTTTCCAC 660
Db      3414 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACTTTTGAGGAAGTGCCTTTCCAC 3473
QY      661 AGCAGCTACGGGCACAGCGCAGAGCTCGGCTGATGAATCCCTCTCATCGACCAATAC 720
```


Db 3826 TTTGGAGCAAGGCTCAGAGAAAACAAATGTGGACATTTGAAAAGGTCAATGATTAACAGAC 3885
Qy 1081 GAAGAGGAATTAAGCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
Db 3886 GAAGAGGAATCAGGAAACCAATCCCGTGGCTACGAGCAGATGTTCTGTATCTACC 3945
Qy 1141 AATTTCCAGAGCAGCAGACAGACCTTCGCGCCGGAGATGTGCATGCTATGGAGCATTTA 1200
Db 3946 AACCTCCAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAACACACAAGCGTTCCT 4005
Qy 1201 CCTGGCATGGTGGCAAGATAGAGACGTGTACTGCAGGGTCCCAATTTGGGCCAAAATTT 1260
Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACTTTCAAGGGGCCCACTCTGGGCAAGATT 4065
Qy 1261 CCTCACACAGATGGACATTTTCACCGCTCTCTCTTTATGGCGGCTTTTGGACTCAAGAAC 1320
Db 4066 CCACACACGAGCGGACATTTTACCCCTCTCCCTCATGGTGGATTCGGACTTAACAC 4125
Qy 1321 CGGCTCTCAGATCCTCATCAAAAACAGCCCTGTTCCTGCGAATCCTCCGGCGGAGTTT 1380
Db 4126 CTTCTCTCCACAGATTTCTCATCAAGAACACCCCGGTACTTCCGAATCCTTCGACCACCTTC 4185
Qy 1381 TCAGCTACAAAGTTGCTTCATTCATCACCAATACTCCACAGGACAAAGTGTGGAA 1440
Db 4186 AGTGGCGCAAGTTTGGCTTCTTCATCACACAGTACTCCACGGGACAGGTTCAGCGTGGAG 4245
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCGCTGGAAATCCCGAAGTGCAGTACACA 1500
Db 4246 ATCGAGTGGAGCTGCAGAAAGGAAAACAGAAACGCTGGAAATCCCGAATTCAGTACACT 4305
Qy 1501 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATACT 1560
Db 4306 TCCAACTACAAACAGTCTGTAATGTGGACTTTACTGTGGACACTAATGGCGGTATTCA 4365
Qy 1561 GAGCCTCGCCCATTTGGACCCGTTACTTACCCGTCCTCTGTA 1604
Db 4366 GAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTGTA 4409

RESULT 9

US-09-770-315-3

; Sequence 3, Application US/09770315

; Patent No. 6429001

; GENERAL INFORMATION:

; APPLICANT: Chiron Corporation

; TITLE OF INVENTION: Recombinant AAV Packaging Systems

; FILE REFERENCE: 20263-501

; CURRENT APPLICATION NUMBER: US/09/770,315

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: US 60/178,536

; PRIOR FILING DATE: 2000-01-26

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 7557

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: recombinant DNA

US-09-770-315-3

Query Match

Best Local Similarity 64.0%; Score 1026.4; DB 3; Length 7557;

Matches 1250; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

Qy 1 ATGGCTTCAGCGGTGGCGCAACCAATGGCAGACAATAACGAAGCGCGCCGAGGTGGGT 60
Db 2837 ATGGCTACAGCAGTGGCGCAACCAATGGCAGACAATAACGAAGCGCGCGAGTGGGT 2896
Qy 61 AATGCTCTCAGGAATTTGGCAATTCGATTCACATGGCTGGCGACAGATCATCACACC 120
Db 2897 AATTCCTCGGGAATTTGGCAATTCGATTCACATGGATGGCGCACAGATCATCACACC 2956

Db 4034 CCAGGATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4093
Qy 1261 CCTCACAGATGGACACTTTACCGGTCTCTCTTATGGGCGGCTTGGACTCAAGAAC 1320
Db 4094 CCACACAGGACGGACATTTTACCCCTCTCCCTCATGGGTGGATTGGACTTAAACAC 4153
Qy 1321 CGCGCTCTCAGATCTCATCAAAACACGCTGTTCCTGCGAATCCCTCCGGCGGAGTTT 1380
Db 4154 CTCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTTGCATCTCTTGACCACTTC 4213
Qy 1381 TCAGCTACAAAGTTTGCTTTTATTCATCACCCCAATACTCCACAGGACAAGTGGTGGAA 1440
Db 4214 AGTGGCGCAAGTTTGCTTCTCTTCATCACAGACTACTCCACGGACAGGTTCAGCGTGGAG 4273
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 4274 ATCGAGTGGAGCTGCAGAAAGAAACAGCAACGCTGGAATCCCGAATTCAGTACACT 4333
Qy 1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATCT 1560
Db 4334 TCAACTACAAAGTCTGTTAATGGACTTTTACTGTGGACACTAATGGCGTGTATTCA 4393
Qy 1561 GAGCCTCGCCCATTTGGCACCCGTTTACCTTACCGTCCCTCTGA 1604
Db 4394 GAGCCTCGCCCATTTGGCACCCAGATACCTGACTCGTAACTCTGA 4437

RESULT 10
US-09-770-315-2
; Sequence 2, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 8698
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-2

Query Match 64.0%; Score 1026.4; DB 3; Length 8698;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1250; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

Qy 1 ATGGCTTACAGGCGGTGGCGACCAATGGCAGACAATAACGAAGGCCCGACAGTGGGT 60
Db 2809 ATGGCTACAGGCGAGTGGCGACCAATGGCAGACAATAACGAAGGCCCGACAGTGGGT 2868
Qy 61 AATGCTTCAGGAAATTTGGCATTCGGATTCCCATGGCTGGCGACAGAGTCATCACACC 120
Db 2869 AATTCTCGGAAATTTGGCATTCGGATTCCCATGGATGGCGACAGAGTCATCACACC 2928
Qy 121 AGCACCGGACCTGGGCGCTTGGCCACTCAATTAACCACTCTACAAAGCAATCTTCAGT 180
Db 2929 AGCACCGGAACCTGGGCGCTTGGCCACTCAATTAACCACTCTACAAAGCAATTTCCAGC 2988
Qy 181 GCTTCAACGGGGCCAGCAACGACCACTACTCTCGGTACAGCACCCCTCTGGGGGTAT 240
Db 2989 CAATCA---GGACCTCGAACGACATCACTATTTGGCTACAGCACCCCTTGGGGGTAT 3045
Qy 241 TTTGATTTCAAGATTTCCACTGCCACTTTTCAACGCTGATGGCAGCGACTCATCAAC 300
Db 3046 TTTGACTTCAACAGATTTCCACTGCCACTTTTCAACGCTGATGGCAGCGACTCATCAAC 3105

Qy 301 AACAAATTGGGATTTCCGGSCCAAGAGACTCAACTTTCAAACCTTTCAACATCAAGTCAAG 360
Db 3106 AACAACTGGGATTTCCGACCCCAAGAGACTCAACTTTCAAGCTTTTAAATTTCAAGTCAAA 3165
Qy 361 GAGGTACAGCAAGATGATGGCGTCAACACCATCGCTATAAACCCTTTACAGCACGGTTCAA 420
Db 3166 GAGGTACAGCAAGATGACGGTACGACGAGATTGGCAATAACCTTTACAGCACGGTTCA 3225
Qy 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCCTCGGCTTCGGCATCGAGGCTGC 480
Db 3226 GTGTTTACTGACTCGGAGTACAGCTTCGGTACGTCCTCGGCTTCGGCATCAAGGATGC 3285
Qy 481 CTCCCTCCGTTCCCGGCGGAGGTTCATGATTCGGCAATACGGCTACTCGACGCTCAAC 540
Db 3286 CTCCCGCGTTCCCGACGACAGCTTCATGTGGCCACAGTATGGATACCTCACCTCGAAC 3345
Qy 541 AATGCAAGCCAGCGTGGGACGTTTCATCTCTTTTACTCGCTTGGAAATATTTCCCTTCTCAG 600
Db 3346 AACGGAGTCAAGGAGTAGGACGCTCTTCATTTTACTCGCTGGAGTACTTTCTTCTCAG 3405
Qy 601 ATGCTGAGAACGGGCAACAACTTTTACCTTACGCTACACCTTTGAGGAAAGTGCCTTTCAC 660
Db 3406 ATGCTGCGTACCGGAAACAACTTTTACCTTACGCTACACCTTTTGGAGGACGTTCTCTTCCAC 3465
Qy 661 AGCAGTACGCGCACAGCCAGAGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 3466 AGCAGTACGCTCACAGCCAGAGCTGGACCGTCTCATGAATCCTCTCATCGACCAATAC 3525
Qy 721 CTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGGCCCAAAACAAGGACTTGTG 780
Db 3526 CTGTATTACTTGAGCAGAACAAACACACTCAAGTGGAAACCAACACGACGCTCAAGGCTTCAG 3585
Qy 781 TTTAGCCGTGGTCTCCAGCTGGGATGCTGTTCAGCCCAAAACAACTGGCTACTGTGACCC 840
Db 3586 TTTTCTCAGGCGCGAGCGAGTGACATTCGGGACCACTAGGAACTGGCTCTCTGGACCC 3645
Qy 841 TGTATTCGGCAGCAGCGGCTTTCTTAAACAAAAACAGACAAACAAACAGCAATTTTACC 900
Db 3646 TGTATTCGGCAGCAGCGAGTATCAAGACATCTCGGATTAACAAACAGTGAATACTCG 3705
Qy 901 TGGACTGGTCTTCAAAATATAAACCCTCAATGGCGTGAATCCATCATCAACCCCTGGCACT 960
Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGTGTAATCCGGGCCG 3765
Qy 961 GCTATGGCTCACAAAGACGACGAAAGATTTCTTCCCATGAGCGGTGTCTATGATT 1020
Db 3766 GCCATGGCAAGCCACAAGACGATGAAGAAAGTTTTTTTCTCTCAGAGCGGGTTCTCATC 3825
Qy 1021 TTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCATTTGGCAATGTCTATGATTACAGAC 1080
Db 3826 TTTTGGAAAGCAAGGCTCAGAGAAAACAAATGTGGACATTTGAAGAGTATGATTATACAGAC 3885
Qy 1081 GAAGAGGAAATTTAAAGCCACTAACCCCTGTGGCACCCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 3886 GAAGAGGAAATCAGGACAAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945
Qy 1141 AATTTCCAGAGCAGCAGACACGCTTCGGACCGGAGATGTGCATGCTATGGGAGCATTA 1200
Db 3946 AACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAAGCGGTTCTT 4005
Qy 1201 CCTGCATGGTGTGGCAAGATAGACGTTGTACTGTACCTGACAGGTCCTATTTGGGCCAAAAT 1260
Db 4006 CAGGAGTGGTGTGGCAGGACAGATGTGTACCTTCAAGGCCCCCATCTGGGCAAGATTT 4065
Qy 1261 CCTCACAGATGAGACACTTTTACCCGCTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320
Db 4066 CCACACAGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCGACTTAAACAC 4125
Qy 1321 CGGCTCTCAGATCCTCATCAAAACACGCGCTGTCTCTGCGAATCTCTCCGCGGAGTTT 1380
Db 4126 CCTCTCTCAACAGATTTCTATCAAGAACACCCCGGTAACCTGCGAATCTCTTCGACCAACCTTC 4185

QY 1381 TCAGCTACAAAGTTTGTCTTCATTCATCACCACCAATCTCCACAGGACA-AGTGAGTGTGGA 1439
DB 4186 AGTGGCGGCAAGTTTGCTTCTTCATCACAACAGTACTCCACGGACACGGTCAGCGTGA 4245
QY 1440 AATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTCAGTACAC 1499
DB 4246 GATCGAGTGGAGCTGCAGAAAGAAAACAGCAACGCTGGAATCCCGAATTCAGTACAC 4305
QY 1500 ATCCAAATTGCAAAATGCCAAGCTTGATTTTACTGTGGCAACAATGGACTTTATAC 1559
DB 4306 TTCCAACTCAACAAGCTGTGTAATCGTGACTTACCGTGGATACTAATGGCGGTGATTC 4365
QY 1560 TGAGCTCGCCCAATGGCAACCGTTACCTTACCTCCCGTCCCTGTA 1604
DB 4366 AGAGCCTCGCCCAATGGCAACAGATACCTGACTCGTAATCTGTA 4410

RESULT 13
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786211and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-475-391-1

Query Match 62.6%; Score 1004.2; DB 2; Length 4680;
Best Local Similarity 77.4%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 358; Indels 4; Gaps 2;
QY 1 ATGGCTTCAGCGGTGGCGCACCAATGCGAGACAATAACGAAGCGCGGACGAGTGGGT 60
DB 2809 ATGGCTACAGCGAGTGGCGCACCAATGCGAGACAATAACGAAGCGCGGACGAGTGGGT 2868
QY 61 AATGCGCTCAGGAATTTGGCAATTGGCAATTGGCAATGGCTGGCGGACAGAGTCATCACCACC 120

DB 2869 AATTCCTCGGAAATTTGGCAATTGGCAATTGCCATGATGGGCGACAGAGTCAATCACCACC 2928
QY 121 AGCACCCGACCTGGGCTCTTGCCCACTACATAAACACACCTCTACAAGCAAAATCTCCAGT 180
DB 2929 AGCACCCGAACTGGGCTCTTGCCCACTACATAAACACACCTCTACAAGCAAAATTTCCAGC 2988
QY 181 GCTTCAAACGGGGGTCAGAAACGACCACTACTCTTCGGGTACAGCAACCCCTTGGGGGTAT 240
DB 2989 CAATCA--GGAGCCTCGAAACGACCAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3045
QY 241 TTTGATTTCAACAGATTTCACTGCGCACCTTTTCCACACGCTGACTGCGAGGAGCTCATCAAC 300
DB 3046 TTTGACTTCAACAGATTTCACTGCGCACCTTTTCCACACGCTGACTGCGAGGAGCTCATCAAC 3105
QY 301 AACAAATTGGGATTTCCGGCCCAAGAGAGACTCAACTTTCAAACTCTTCAACATCCAAGTCAAG 360
DB 3106 AACAACTGGGATTTCCGACCCCAAGAGAGCTCAACTTTCAAGCTCTTTAAACATTTCAAGTCAAA 3165
QY 361 GAGGTACAGCAAGATGATGGCGTCAACACCATCGCTTAATAAATTTACAGAGAGCGTTCAA 420
DB 3166 GAGGTACAGCAAGATGATGGCGTCAACACCATCGCTTAATAAATTTACAGAGAGCGTTCAA 3225
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGCGCACACAGGCGTGC 480
DB 3226 GTGTTTACTGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGCGCATCAAGATGC 3285
QY 481 CTCCCTTCGGTTCCTCGGCGGACGCTTTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 540
DB 3286 CTCCCGCGTTCCTCGGCGGACGCTTTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 3345
QY 541 AATGGCAGCCAAAGCGTGGAGCGTTCACTCTTTTACTGCTGGAATATTTTCCCTTCTCAG 600
DB 3346 AACGGGAGTCAGGCAAGTAGGAGCGCTCTTCATTTTACTGCTGGAGTACTTCTCTTCTCAG 3405
QY 601 ATGCTGAGAACGGGCAACAACTTTTACTTACCTACCTACCTTTGAGGAAGTGCCTTTCCAC 660
DB 3406 ATGCTGCGTACCGGAACAACTTTTACTTACCTACCTTTTGGAGGAGCTTCTCTTCCAC 3465
QY 661 AGCAGCTACCGCACAGCGCAGAGCTGGACCGGCTGATGAATCCTCTCATCGACCAATATC 720
DB 3466 AGCAGCTACCGTCAACAGCGAGAGTCTGGACCGCTCTCATGAATCCTCTCATCGACCAATATC 3525
QY 721 CTGTATTACTTGAACAGAACTCAAAATCAGTCCGGAAGTGCACCAACAAAGGACTTGTGTG 780
DB 3526 CTGTATTACTTGAACAGAACTCAAAATCAGTCCGGAAGTGCACCAACAAAGGACTTGTGTG 3585
QY 781 TTTAGCGGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAACTGCTGCTGACCTGACCC 840
DB 3586 TTTTCTCAGGCGGAGCGAGTGACATTCGGGACCACTCTAGGAACCTGGCTTCTCTGACCC 3645
QY 841 TGTATTGCGCAGCAGCGCGCTTTCTAAACAAAAACAGACAAACAAACAGCAATTTTACC 900
DB 3646 TGTATTGCGCAGCAGCGAGTATCAAGACATCTGCGGATACAAACAGAGTGNATCTCG 3705
QY 901 TGGACTGTGTCTTCAAAATATATAACCTCAATGGGCGTGAATCCATCATCAACCTCGGACT 960
DB 3706 TGGACTGTGAGCTACCAAGTACCACTCAATGGCAGAGACTCTCTGTGTGAATCCGGGCGCC 3765
QY 961 GCTATGGGCTCACAACAGACGAGAGCAAGTCTTTCCCATGAGCGGTGTGATGATT 1020
DB 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCTCAGAGCGGGTCTCATC 3825
QY 1021 TTTGGAAGAGAGAGCGCGGAGCTTCAAAACACTGCAATTTGGAATGTGATGATTACAGAC 1080
DB 3826 TTTGGAAGAGAGCTCAGAGAAAACAAATGTGAACATTTGAAGAGTCAATGATTACAGAC 3885
QY 1081 GAAGAGGAATTAAGGCCACTAAACCTGTGGCCACCGAAAGATTTTGGGACCGTGGCAGTC 1140
DB 3886 GAAGAGGAATCGGAACAACTCAATCCGTGGCTACGGAGCAGTATGTTTCTGTATCTACC 3945
QY 1141 AATTTCCAGAGCAGCAGCAGACCTCGGAGATGTGATGCTATGGGAGCATTA 1200
DB 3946 AACCTCCAGAGGCAACAGACGAGCTACCGCAGATGTCAACACACAAAGCGCTTCTT 4005

QY 1201 CCTGGCATGGTGGCAAGATAGACGCGTGACCGAGGTCCTCCATTTGGGCCAAATTT 1260
DB 4006 CAGGGCATGGTCTGGCAGACAGAGATGTGTACCTTCAGGGGCCCAATCTGGGCAAGATT 4065
QY 1261 CCTCACACAGATGGACACTTTTCCACCGCTCTCTCTTATATGGGGGGCTTTGGACTCAAGAAC 1320
DB 4066 CCACACACGACGGACATTTTCCACCCCTCTCCCTCATGGGTGGATTGGACTTAAACAC 4125
QY 1321 CGGCTCTCAGATCTCATCAAAACACCGCTGTTCCTTGGGAAATCTCCGGCGGAGTTT 1380
DB 4126 CCTCTCTCCACAGATTTCTCATCAAGAACACCCCGGTACCTTGGGAAATCTTTCGACCACCTTC 4185
QY 1381 TCAGCTACAAAGTTTGTCTTCAATTCATCCCAATATCTCCACAGGACA-AGTCAGGTGGA 1439
DB 4186 AGTCGGGCAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGACACGGTCAGCGTGA 4245
QY 1440 AATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACAC 1499
DB 4246 GATCGAGTGGAGCTGCAGAGGAAAACAGCAACGCTGGAAATCCCGAAATTCAGTACAC 4305
QY 1500 ATCCAAATTAGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAAATGGACTTTATAC 1559
DB 4306 TTCCAACTACAAAGTCTGTAAATCGTGGACTTACCGTGGATCTTAATGGCGTGTATTC 4365
QY 1560 TGAGCTCGCCCATTTGGCACCCTGTACCTTACCGTCCCTCGTA 1604
DB 4366 AGAGCCTCGCCCATTTGGCACCAGATACCTGTGTAATCTGTA 4410

RESULT 14

US-08-709-609-1
/ Sequence 1, Application US/08709609
/ Patent No. 5858775
/ GENERAL INFORMATION:
/ APPLICANT: Johnson, Philip R.
/ TITLE OF INVENTION: Adeno-Associated Virus Materials and
/ TITLE OF INVENTION: Methods
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
/ STREET: 6300 Sears Tower, 233 S. Wacker Drive
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60606
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: US/08/709,609
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 5858775and, Greta E.
/ REGISTRATION NUMBER: 35,302
/ REFERENCE/DOCKET NUMBER: 31975
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 474-6300
/ TELEFAX: (312) 474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4680 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-709-609-1
Query Match 62.6%; Score 1004.2; DB 2; Length 4680;

Best Local Similarity 77.4%; Pred. No. 0;
Matches 1243; Conservative 0; Mismatches 358; Indels 4; Gaps 2;
QY 1 ATGGCTTCAGCGGTGGCGACCAATGGCAGACAATAACGAAGCGCGCAGCGAGTGGGT 60
DB 2809 ATGGCTTCAGCGGTGGCGACCAATGGCAGACAATAACGAAGCGCGCAGCGAGTGGGT 2868
QY 61 AATGCTTCAGGAAATTTGGCATTTCCACATGGCTGGGGCGACAGAGTTCATCAACACC 120
DB 2869 AATGCTTCAGGAAATTTGGCATTTCCACATGGATGGGGCGACAGAGTTCATCAACACC 2928
QY 121 AGCACCCGACCTTGGCCCTTCCACACCTTCAATAACCACTCTCAACGAACAATCTCCAGT 180
DB 2929 AGCACCCGAACTTGGCCCTTCCACACCTTCAACCAACCACTCTCAACGAACAATTTCCAGC 2988
QY 181 GCTTCAACGGGGGCGAGCAACGACCACTTCTTGGCTACAGCACCCCTGGGGGTAT 240
DB 2989 CAATCA---GGAGCCTCGAAGCAATCACTTCTTGGCTACAGCACCCCTTGGGGGTAT 3045
QY 241 TTTGATTTCAACAGATTTCCACTGCCACTTTTCAACAGTGTGCGAGGACTCATCAAC 300
DB 3046 TTTGATTTCAACAGATTTCCACTGCCACTTTTCAACAGTGTGCGAGGACTCATCAAC 3105
QY 301 AACAAATGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAGTCAAG 360
DB 3106 AACAACTGGGATTTCCGACCCCAAGAGACTCAACTTTCAAGCTCTTTAACTCAAGTCAAA 3165
QY 361 GAGGTCAACGAAATGATGGCGTCAACCACTTGGCTAATAACCTTTACAGACAGCGTTCAA 420
DB 3166 GAGGTCAACGAAATGATGGCGTCAACCACTTGGCTAATAACCTTTACAGACAGCGTTCAA 3225
QY 421 GTCTTCTCGGACTCGGAGTACAGGCTTCGCTGCTCGGCTCGCCACAGCGGTGC 480
DB 3226 GTGTTTACTGACTCGGAGTACAGGCTTCGCTGCTCGGCTCGCCGATCAAGGATGC 3285
QY 481 CTCCCTCTCGTTCCTCGGCGAGCGTGTTCATGATTTCCGCAATACGGCTACTGACGCTCAAC 540
DB 3286 CTCCCTCGGCTTCCTCGGCGAGCGTGTTCATGATTTCCGCAATACGGCTACTGACGCTCAAC 3345
QY 541 AATGCGACGACGCGTGGAGTTCATCTTTTACTGCTGGAAATATTTTCCCTTCTCAG 600
DB 3346 AACGGGAGTCAAGGAGTGGAGCGCTCTTCAATTTTACTGCTGGAGTACTTCTCTTCTCAG 3405
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCTTTCCAC 660
DB 3406 ATGCTGCGTACCGGAAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCTTTCCAC 3465
QY 661 AGCAGCTACGCGCACAGCAGCGCTGGACCGGCTGATGAATCTCTCTCATCGACCAATAC 720
DB 3466 AGCAGCTACGCTACAGCAGCGCTGGACCGCTCTCATGAATCTCTCTCATCGACCAATAC 3525
QY 721 CTGTATTACTGAAACAGAACTCAAAATCAGTCCGGAAGTGGCCAAACAGAGACTTGTGTG 780
DB 3526 CTGTATTACTTGAGCAGAACAACTCAAGTGGAAACCAACCCAGCAGTCAAGGCTTCAAG 3585
QY 781 TTTAGCGTGGGTCTCGAGCTGGCATGTCTGTTCAGCCCAAAACCTGCTACTCGGACCC 840
DB 3586 TTTTCTCAGGCGCGGAGCGAGTGACATTCGGGACCAAGTCTAGGAATCTGCTTCTGAGCC 3645
QY 841 TGTATTTCGCGCAGCGGGGTTCCTAAACAAAAACAGACAAACAAACAGCAATTTTAC 900
DB 3646 TGTATTTCGCGCAGCGGGATTCAGGATTCAGGATTCAGGATTCAGGATTCAGGATTCAG 3705
QY 901 TCGACTGGGTCTCAAAATATATACCTCAATGGGCTGATTCATCATCAACCTCGGACT 960
DB 3706 TGGACTGGAGCTTACCAAGTACCACTCAATGGCAGAGACTCTCTGCTGAAATCCGGGGCCC 3765
QY 961 GCTATGGCTCTCACAAAGACGACCAAGAGTCTTTTCCCATGAGCGGTGCTCATGATT 1020
DB 3766 GCCATGGCAAGCCCAAGGACGATGAAGAAAAAGTTTTTCTCCTCAGAGCGGGTCTCATC 3825
QY 1021 TTTGAAAGAGAGCGCGGAGCTTTCAACCACTGCAATTTGGCAATGTCTATGATTACAGAC 1080

Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGGCCC 3765
Qy 961 GCTATGGGCTCACAACAAGACGACGAAGATTCTTTCCCATGAGCGGTGTCATGATT 1020
Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTCTCTCAGAGCGGGTTCTCATC 3825
Qy 1021 TTTGGAAAGAGAGCGCGGAGCTTCAACACCTGCATTGGACAATGTGATTACAGAC 1080
Db 3826 TTTGGAAAGCAAGGCTCAGAGAAAACAATGTGAACATTTGAAAAGGTATGATTACAGAC 3885
Qy 1081 GAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 3886 GAAGAGGAAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945
Qy 1141 AATTTCCAGAGCAGCAGCAGACCGCTGGACCGGAGATGTGCATGCTATGGGAGCATTA 1200
Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGCGTTCTT 4005
Qy 1201 CCTGGCATGGTGTGCAAGATAGAGACGTGTACTGTGAGGGTCCCATTTGGGCCAAATTT 1260
Db 4006 CCAGCATGGTCTGGCAGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065
Qy 1261 CCTCACACAGATGGACACTTTTCAACCGCTCTCCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320
Db 4066 CCACACAGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTGGACTTAAACAC 4125
Qy 1321 CCGCCTCCTCAGATCCTCATCAAAAACAGCGCTGTTCCTGCGAATCCTCCGGCGGAGTTT 1380
Db 4126 CTTCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 4185
Qy 1381 TCAGCTACAAAGTTTGTCTTTCATTTCATCACCCTCAATCTCCACAGGACA-AGTGAGTGTGA 1439
Db 4186 AGTGGGCAAGTTTGTCTTCTTCATCACAGTACTCCACGGGACACGGTCAGCGTGA 4245
Qy 1440 AATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACAC 1499
Db 4246 GATCGAGTGGAGCTGCAGAGGAAAACAGCAACCGCTGGAATCCCGAATTCAGTACAC 4305
Qy 1500 ATCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTTTATAC 1559
Db 4306 TTCCAACCTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGGTGATTTC 4365
Qy 1560 TGAGCCTCGCCCATTTGGCACCCGTTTACCCTTACCCGTCCTCTGTA 1604
Db 4366 AGAGCCTCGCCCATTTGGCACCAAGATACCTGACTCGTAATCTGTA 4410

Search completed: November 29, 2005, 14:27:50
Job time : 233.137 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 08:46:14 ; Search time 750.542 Seconds
(without alignments)
14243.261 Million cell updates/sec

Title: US-10-696-900-1_COPY_2829_4432
Perfect score: 1604
Sequence: 1 atgggttcaggcggtgggc.....tacctaccgtccctgtga 1604

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005as.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1604	100.0	1605	3 AAD00779	Ad00779 Adeno-ss
2	1604	100.0	1800	3 AAD00778	Ad00778 Adeno-ss
3	1604	100.0	2211	3 AAD00777	Ad00777 Adeno-ss
4	1604	100.0	2211	14 AD227052	Ad227052 Adeno-ss
5	1604	100.0	2211	14 AD226929	Ad226929 Adeno-ss
6	1604	100.0	4347	13 ADW39398	Adw39398 Adeno-ss
7	1604	100.0	4347	14 AD246594	Ad246594 HSV-AAV s
8	1604	100.0	4718	3 AAD00772	Ad00772 Adeno-ss
9	1604	100.0	4718	10 AD76507	Ad76507 Adeno-ss
10	1604	100.0	4718	10 ADL13984	Adl13984 Adeno-ss
11	1604	100.0	4718	12 ADG39758	Adg39758 AAV-1 gen
12	1604	100.0	7447	8 ACC58477	Acc58477 Vector pf
13	1600.8	99.8	2211	14 AD226931	Ad226931 Adeno-ss
14	1594.4	99.4	2214	14 AD226930	Ad226930 Adeno-ss
15	1592.8	99.3	2211	14 AD226928	Ad226928 Adeno-ss
16	1591.2	99.2	2211	14 AD226932	Ad226932 Adeno-ss
17	1512.8	94.3	2211	14 AD227053	Ad227053 Adeno-ss
18	1512.8	94.3	4239	13 ADW39402	Adw39402 Adeno-ss
19	1512.8	94.3	4239	14 AD246598	Ad246598 HSV-AAV s

20	1512.8	94.3	4683	4 AAF23749	Aaf23749 AAV6 DNA
21	1512.8	94.3	4683	10 ADL13983	Adl13983 Adeno-ss
22	1512.8	94.3	4683	12 ADG39763	Adg39763 AAV-6 gen
23	1104.6	68.9	2214	14 ADZ26890	Adz26890 Adeno-ss
24	1084	67.6	2214	14 ADZ26891	Adz26891 Adeno-ss
25	1079	67.3	2214	14 ADZ26886	Adz26886 Adeno-ss
26	1077.4	67.2	2214	14 ADZ26885	Adz26885 Adeno-ss
27	1075.8	67.1	2217	14 ADZ26892	Adz26892 Adeno-ss
28	1074.2	67.0	2214	14 ADZ26884	Adz26884 Adeno-ss
29	1074.2	67.0	2214	14 ADZ26882	Adz26882 Adeno-ss
30	1072.6	66.9	2214	14 ADZ26889	Adz26889 Adeno-ss
31	1071.6	66.8	3094	10 ADE76523	Ad76523 Adeno-ss
32	1071	66.8	2214	14 ADZ26883	Adz26883 Adeno-ss
33	1070	66.7	2214	14 ADZ26887	Adz26887 Adeno-ss
34	1070	66.7	2214	14 ADZ27063	Adz27063 Adeno-ss
c 35	1070	66.7	3095	10 ADE76525	Ad76525 Adeno-ss
c 36	1070	66.7	3095	10 ADE76524	Ad76524 Adeno-ss
37	1070	66.7	4721	10 ADE76502	Ad76502 Adeno-ss
38	1070	66.7	4721	12 ADG39764	Adg39764 AAV-7 gen
39	1070	66.7	4721	14 ADV67509	Adv67509 Nucleotid
40	1068.4	66.6	3127	10 ADZ27030	Adz27030 Adeno-ss
c 41	1068.4	66.6	3127	10 ADE76516	Ad76516 Adeno-ss
42	1068	66.6	2209	14 ADZ26898	Adz26898 Adeno-ss
43	1068	66.6	3075	10 ADE76527	Ad76527 Adeno-ss
44	1066.4	66.5	2208	14 ADZ26895	Adz26895 Adeno-ss
45	1064.8	66.4	2208	14 ADZ26897	Adz26897 Adeno-ss

ALIGNMENTS

RESULT 1

AAD00779	
ID	AAD00779 standard; DNA; 1605 BP.
XX	
AC	AAD00779;
XX	
DT	08-SEP-2000 (first entry)
XX	
DE	Adeno-associated virus serotype 1 capsid protein VP3 DNA.
XX	
KW	Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW	cap protein; recombinant viral vector; gene delivery; gene therapy;
KW	vaccine; transgene; VP3; ss.
XX	
OS	Adeno associated virus serotype 1.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1605
FT	/*tag= a
FT	/product= "VP3 protein"
XX	
PN	WO200028061-A2.
XX	
PD	18-MAY-2000.
XX	
PF	02-NOV-1999; 99WO-US025694.
XX	
PR	05-NOV-1998; 98US-0107114P.
XX	
PA	(UYPE-) UNIV PENNSYLVANIA.
XX	
PI	Wilson JM, Xiao W;
XX	
DR	WPI; 2000-376571/32.
DR	P-PSDB; AAY71169.
XX	
PT	Novel adeno-associated virus serotype 1 polynucleotide useful for
PT	preparation of medicament for delivery of a transgene to a host.
XX	
PS	Claim 10; Page 96-99; 108pp; English.
XX	
CC	The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA

which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions, are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV hence is useful for transforming host cells, and in the preparation of a medicament for the delivery of transgene to a host. The present sequence is an AAV-1 DNA encoding a cap protein VP3 which is useful in the production of recombinant viral vector for gene delivery

XX Sequence 1605 BP; 431 A; 466 C; 361 G; 347 T; 0 U; 0 Other;

Query Match		100.0%;	Score 1604;	DB 3;	Length 1605;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1604;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCTTCAGGGCGTGGCGCAACAATGGCAGACAATAACGAAGCGCCGACGGAGTGGGT	60		
Db	1	ATGGCTTCAGGGCGTGGCGCAACAATGGCAGACAATAACGAAGCGCCGACGGAGTGGGT	60		
QY	61	AATGCTCAGGAATTTGGCATTTGGGATTCACATGGCTGGGCGACAGAGTCATCACCAC	120		
Db	61	AATGCTCAGGAATTTGGCATTTGGGATTCACATGGCTGGGCGACAGAGTCATCACCAC	120		
QY	121	AGCACCGCACCTGGGCGCTTGGCCACCTACCAATACCACTCTTACAGCAAAATCTCCAGT	180		
Db	121	AGCACCGCACCTGGGCGCTTGGCCACCTACCAATACCACTCTTACAGCAAAATCTCCAGT	180		
QY	181	GCTTCAACGGGGCCAGCAACGACAAACCACTACTTCGGGTACAGCAACCCCTCGGGGGTAT	240		
Db	181	GCTTCAACGGGGCCAGCAACGACAAACCACTACTTCGGGTACAGCAACCCCTCGGGGGTAT	240		
QY	241	TTTGATTTCAACAGATTCACATGGCCACTTTTACCAACGATGACTGGCAGGACTCATCAAC	300		
Db	241	TTTGATTTCAACAGATTCACATGGCCACTTTTACCAACGATGACTGGCAGGACTCATCAAC	300		
QY	301	AACAAATGGGGATTCGGGCGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG	360		
Db	301	AACAAATGGGGATTCGGGCGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG	360		
QY	361	GAGGTCAAGCAAGATGATGGGTCAACAACATGCTTAATAACCTTACAGACAGCGTTCAA	420		
Db	361	GAGGTCAAGCAAGATGATGGGTCAACAACATGCTTAATAACCTTACAGACAGCGTTCAA	420		
QY	421	GTCTCTCGGACTCGGAGTACAGCTTCGGTACGTTCCGCTCTCGGCTCGGCGACAGGCTGC	480		
Db	421	GTCTCTCGGACTCGGAGTACAGCTTCGGTACGTTCCGCTCTCGGCTCGGCGACAGGCTGC	480		
QY	481	CTCCCTCCGTTCCCGCGGACAGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC	540		
Db	481	CTCCCTCCGTTCCCGCGGACAGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC	540		
QY	541	AATGGCAGCAAGCGTGGGACGTTTCATCTTTTACTGCTGGAAATTTTCCCTTCTCAG	600		
Db	541	AATGGCAGCAAGCGTGGGACGTTTCATCTTTTACTGCTGGAAATTTTCCCTTCTCAG	600		
QY	601	ATGCTGAGACGGGCAACACTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC	660		
Db	601	ATGCTGAGACGGGCAACACTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC	660		
QY	661	AGCAGCTACGCGCACAGCGCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC	720		
Db	661	AGCAGCTACGCGCACAGCGCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC	720		
QY	721	CTGTATTACCTGAACAGAACTCAAAATCAAGTCCGGAAGTGCCTCAAAACAGGACTTGTG	780		
Db	721	CTGTATTACCTGAACAGAACTCAAAATCAAGTCCGGAAGTGCCTCAAAACAGGACTTGTG	780		
QY	781	TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACTGGCTACCTGGACCC	840		

Db	781	TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACTGGCTACCTGGACCC	840
QY	841	TGTTATCGGACGAGCGGTTTCTTAAACAACAAAACAGACAACAACAGCAATTTTACC	900
Db	841	TGTTATCGGACGAGCGGTTTCTTAAACAACAAAACAGACAACAACAGCAATTTTACC	900
QY	901	TGGACTGTGCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT	960
Db	901	TGGACTGTGCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT	960
QY	961	GCTATGGCTCACAACAAGACGACGAAGAAGTCTTCTTCCATGAGCGGTGTCATGATT	1020
Db	961	GCTATGGCTCACAACAAGACGACGAAGAAGTCTTCTTCCATGAGCGGTGTCATGATT	1020
QY	1021	TTTGGAAAAGAGAGCGCGGAGCTTCAAACTGCAATGGAGCAATGTGTCATGATTACAGAC	1080
Db	1021	TTTGGAAAAGAGAGCGCGGAGCTTCAAACTGCAATGGAGCAATGTGTCATGATTACAGAC	1080
QY	1081	GAAGAGGAATTTAAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140
Db	1081	GAAGAGGAATTTAAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140
QY	1141	AATTTCCAGAGCAGCAGCACAGACCCCTCGGACCGGAGATGTGCATGTATGGGACATTAT	1200
Db	1141	AATTTCCAGAGCAGCAGCACAGACCCCTCGGACCGGAGATGTGCATGTATGGGACATTAT	1200
QY	1201	CTTGCGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCCAAAATT	1260
Db	1201	CTTGCGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCCAAAATT	1260
QY	1261	CCTCACAGATGGACACTTTTCAACCGCTCTCTTATGGGCGGCTTTGGACTCAAGAAC	1320
Db	1261	CCTCACAGATGGACACTTTTCAACCGCTCTCTTATGGGCGGCTTTGGACTCAAGAAC	1320
QY	1321	CGCGCTCTCAGATCCTCATCAAAAACACGCTGTCTCTGCGAATCTCTCGGCGGAGTTT	1380
Db	1321	CGCGCTCTCAGATCCTCATCAAAAACACGCTGTCTCTGCGAATCTCTCGGCGGAGTTT	1380
QY	1381	TCAGCTCAAAAGTTTGTCTTCAATCATCCCAATACTCCACAGGACAAAGTGTGGAA	1440
Db	1381	TCAGCTCAAAAGTTTGTCTTCAATCATCCCAATACTCCACAGGACAAAGTGTGGAA	1440
QY	1441	ATTGAATGGAGCTCAGAAAAGAAAACAGACGCTGGAAATCCCGAAGTGCAGTACACA	1500
Db	1441	ATTGAATGGAGCTCAGAAAAGAAAACAGACGCTGGAAATCCCGAAGTGCAGTACACA	1500
QY	1501	TCCAATTTATGCAAAATCTGCCAACCTTGAATTTTACTGTGGACAACAATGGACTTTTACT	1560
Db	1501	TCCAATTTATGCAAAATCTGCCAACCTTGAATTTTACTGTGGACAACAATGGACTTTTACT	1560
QY	1561	GAGCCTCGCCCAATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA	1604
Db	1561	GAGCCTCGCCCAATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA	1604

RESULT 2

AAD00778

ID AAD00778 standard; DNA; 1800 BP.

XX AAD00778;

AC AAD00778;

XX 08-SEP-2000 (first entry)

DT 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 capsid protein VP2 DNA.

DE Adeno-associated virus serotype 1: AAV-1; rep protein; capsid protein;

KW cap protein; recombinant viral vector; gene delivery; gene therapy;

KW vaccine; transgene; VP2; ss.

XX Adeno associated virus serotype 1.

OS Adeno associated virus serotype 1.

XX Key

XX Location/Qualifiers

CDS 1..1800
FT FT /*tag= a
FT FT /product= "VP2 protein"
FT FT /partial
XX
PN WO200028061-A2.
PD 18-MAY-2000.
XX 02-NOV-1999; 99WO-US025694.
XX 05-NOV-1998; 98US-0107114P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Wilson JM, Xiao W;
XX WPI; 2000-376571/32.
DR P-PSDB; AAY71168.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host.
XX
XX Claim 10; Page 90-93; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antisense delivery vectors. The AAV-1 does not induce the
CC formation of neutralising antibodies specific to any serotype of AAV
CC hence is useful for transforming host cells, and in the preparation of a
CC medicament for the delivery of transgene to a host. The present sequence
CC is an AAV-1 DNA encoding a cap protein VP2 which is useful in the
CC production of recombinant viral vector for gene delivery
XX
SQ Sequence 1800 BP; 483 A; 532 C; 409 G; 376 T; 0 U; 0 Other;
Query Match 100.0%; Score 1604; DB 3; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGCGCGTGGCGCACAATGGCGAGACAATAACGAAGCGCGCGAGTGGGT 60
DB 196 ATGGCTTCAGCGCGTGGCGCACAATGGCGAGACAATAACGAAGCGCGCGAGTGGGT 255
QY 61 AATGCCCTCAGGAATTTGGCAATTGGGATTCCACATGGCTGGCGACAGAGTCAACACC 120
DB 256 AATGCCCTCAGGAATTTGGCAATTGGGATTCCACATGGCTGGCGACAGAGTCAACACC 315
QY 121 AGCACCGCACCTGGGCTTGGCCACCTCAATAAACCACTCTACAGCAAAATCTCCAGT 180
DB 316 AGCACCGCACCTGGGCTTGGCCACCTCAATAAACCACTCTACAGCAAAATCTCCAGT 375
QY 181 GCTTCAACGGGGCCAGCAACAGCAACTACTCTGGGCTACAGCACCCCTGGGGGTAT 240
DB 376 GCTTCAACGGGGCCAGCAACAGCAACTACTCTGGGCTACAGCACCCCTGGGGGTAT 435
QY 241 TTTGATTTCAACAGATTCACCTGCCATTTTCCACAGTGACTGGGAGGACTCATCAAC 300
DB 436 TTTGATTTCAACAGATTCACCTGCCATTTTCCACAGTGACTGGGAGGACTCATCAAC 495
QY 301 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360
DB 496 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 555
QY 361 GAGGTACGACGAATGATGGCGTCACAAACCATCGCTAATTAACCTTACGACAGCGTTCAA 420
DB 556 GAGGTACGACGAATGATGGCGTCACAAACCATCGCTAATTAACCTTACGACAGCGTTCAA 615

QY 421 GTCTTCTCGGACTCGGAGTACCAAGTTCGGTACGTCTCGGCTCTGCGCACAGGGCTGC 480
DB 616 GTCTTCTCGGACTCGGAGTACCAAGTTCGGTACGTCTCGGCTCTGCGCACAGGGCTGC 675
QY 481 CTCCTCTCGGTTCCCGCGGACGTTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 540
DB 676 CTCCTCTCGGTTCCCGCGGACGTTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 735
QY 541 AATGCGAGCAAGCGTGGGACGTTTCATCTTTTACTGCTTGGAAATATTTCCTTCTCAG 600
DB 736 AATGCGAGCAAGCGTGGGACGTTTCATCTTTTACTGCTTGGAAATATTTCCTTCTCAG 795
QY 601 ATGCTGAGAAACGGGCAACAACTTTTACCTTTCAGTACCTACCTTTGAGGAAGTGCCTTCCAC 660
DB 796 ATGCTGAGAAACGGGCAACAACTTTTACCTTTCAGTACCTACCTTTGAGGAAGTGCCTTCCAC 855
QY 661 AGCAGTACGCGCACAGCGGACGCTGGACCGCTGTAATCCTCTCATCGACCAATAC 720
DB 856 AGCAGTACGCGCACAGCGGACGCTGGACCGCTGTAATCCTCTCATCGACCAATAC 915
QY 721 CTGTATTACCTCAACAGAACTCAAATCAGTCCGGAAGTCCCAAAACAAAGGACTTGTCTG 780
DB 916 CTGTATTACCTCAACAGAACTCAAATCAGTCCGGAAGTCCCAAAACAAAGGACTTGTCTG 975
QY 781 TTTAGCCGTGGTCTCCAGTGGCATGTCTGTTCAGCCCAAAACTGGCTACCTGGACCC 840
DB 976 TTTAGCCGTGGTCTCCAGTGGCATGTCTGTTCAGCCCAAAACTGGCTACCTGGACCC 1035
QY 841 TGTATTCCGACAGCGCGTTTCAAAACAAACAGCAACAAACAAACAGCAATTTTACC 900
DB 1036 TGTATTCCGACAGCGCGTTTCAAAACAAACAGCAACAAACAAACAGCAATTTTACC 1095
QY 901 TGGACTGTGTCTTAAATATAAATTAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960
DB 1096 TGGACTGTGTCTTAAATATAAATTAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 1155
QY 961 GCTATGGCTCTCACAAACAGCAAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1020
DB 1156 GCTATGGCTCTCACAAACAGCAAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1215
QY 1021 TTTGAAAAGAGAGCGCCGAGCTTCAACACTGCAATTCGCAATTCGCAATTCATGATTACAGAC 1080
DB 1216 TTTGAAAAGAGAGCGCCGAGCTTCAACACTGCAATTCGCAATTCGCAATTCATGATTACAGAC 1275
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTCGGACCGTGGCAGTC 1140
DB 1276 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTCGGACCGTGGCAGTC 1335
QY 1141 AATTTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCAATGAGGAGCATTA 1200
DB 1336 AATTTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCAATGAGGAGCATTA 1395
QY 1201 CTTGGCATGGTGGCAAGATAGAGAGCTGTACCTGCGGCTCCCATTTGGGCGCAAAAT 1260
DB 1396 CTTGGCATGGTGGCAAGATAGAGAGCTGTACCTGCGGCTCCCATTTGGGCGCAAAAT 1455
QY 1261 CCTCACACAGATGGACACTTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
DB 1456 CCTCACACAGATGGACACTTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1515
QY 1321 CCGCTCTCTCAGATCTCTCATCAAAAACACGCTGTTCCTGCGAATCCTCGGCGGAGTTT 1380
DB 1516 CCGCTCTCTCAGATCTCTCATCAAAAACACGCTGTTCCTGCGAATCCTCGGCGGAGTTT 1575
QY 1381 TCAGCTACAAAGTTTGGTTCATTCATCACCAATACTCCACAGGACAAAGTGTGGAA 1440
DB 1576 TCAGCTACAAAGTTTGGTTCATTCATCACCAATACTCCACAGGACAAAGTGTGGAA 1635
QY 1441 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGTGGAAATCCCGAAGTGCAGTACACA 1500
DB 1636 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGTGGAAATCCCGAAGTGCAGTACACA 1695

1141 AATTTCCAGAGCAGCAGACACCTCGACCGGAGATGTGCATGCTATGGGACATTA 1200
1747 AATTTCCAGAGCAGCAGCAGACCTCGACCGGAGATGTGCATGCTATGGGACATTA 1806
1201 CTTGCATGTTGTGCAAGATAGAGACGTTACCTGTCAGGGTCCATTTGGGCCAAAATT 1260
1807 CTTGCATGTTGTGCAAGATAGAGACGTTACCTGTCAGGGTCCATTTGGGCCAAAATT 1866
1261 CTTGCATGTTGTGCAAGATAGAGACGTTACCTGTCAGGGTCCATTTGGGCCAAAATT 1320
1867 CTTGCATGTTGTGCAAGATAGAGACGTTACCTGTCAGGGTCCATTTGGGCCAAAATT 1926
1321 CCGCTCTCTCAGATCCTCATCAAAAACACGCTGTTCCTGCGAATCCTCCGGCGGAGTTT 1380
1927 CCGCTCTCTCAGATCCTCATCAAAAACACGCTGTTCCTGCGAATCCTCCGGCGGAGTTT 1986
1381 TCAGCTACAAAGTTTGTCTTTCATTCATCACCCTAATCTCCACAGGACAAAGTGTGGAA 1440
1987 TCAGCTACAAAGTTTGTCTTTCATTCATCACCCTAATCTCCACAGGACAAAGTGTGGAA 2046
1441 ATTGAATGGAGCTCGAAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
2047 ATTGAATGGAGCTCGAAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 2106
1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGCAACAATGGACTTTTATCT 1560
2107 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGCAACAATGGACTTTTATCT 2166
1561 GAGCTCGCCCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGA 1604
2167 GAGCTCGCCCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGA 2210

RESULT 4

AD227052
ID AD227052 standard; DNA; 2211 BP.
XX
AC AD227052;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus DNA SEQ ID NO 202.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antipsoaritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoaritic; vasotropic; gastrointestinal-gen; hemostatic; anti-HIV;
KW viricide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
PN WO200503321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004WO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
XX
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, psoriasis, scleroderma, hemophilia, HIV, bacterial
XX infection and cancer.

PS Disclosure; SEQ ID NO 202; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;
Query Match 100.0%; Score 1604; DB 14; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACATAACGAAGGCGCGAGTGGGT 60
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACATAACGAAGGCGCGAGTGGGT 666
QY 61 AATGCTTCAGGAAATTTGGCATTTCCATGCTGGTGGCGCAGAGTCAATCACCAC 120
Db 667 AATGCTTCAGGAAATTTGGCATTTCCATGCTGGTGGCGCAGAGTCAATCACCAC 726
QY 121 AGCACCCGACCTGGGCTTGGCCACCTACATAACCACTCTCAAGCAAAATCTCCAGT 180
Db 727 AGCACCCGACCTGGGCTTGGCCACCTACATAACCACTCTCAAGCAAAATCTCCAGT 786
QY 181 GCTTCAACGGGGCGCCAGCAACCACTACTCTCGGCTACAGCACCCCTGGGGGTAT 240
Db 787 GCTTCAACGGGGCGCCAGCAACCACTACTCTCGGCTACAGCACCCCTGGGGGTAT 846
QY 241 TTTGATTTCAACAGATTCCATGCGCATTTTCAACAGCTGACTGCGAGGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCCATGCGCATTTTCAACAGCTGACTGCGAGGACTCATCAAC 906
QY 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTCAACATCCAAGTCAAG 360
Db 907 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTCAACATCCAAGTCAAG 966
QY 361 GAGGTCAGCAGCAATGATGGGTCAACCACTCGTAACTTAACAGCAGCGTTCAC 420
Db 967 GAGGTCAGCAGCAATGATGGGTCAACCACTCGTAACTTAACAGCAGCGTTCAC 1026
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGTTACGTTCTCGGCTCTGCGCACAGGCTGC 480
Db 1027 GTCTTCTCGGACTCGGAGTACAGCTTCGTTACGTTCTCGGCTCTGCGCACAGGCTGC 1086
QY 481 CTCCCTTCGGTTCGGCGGAGCGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Db 1087 CTCCCTTCGGTTCGGCGGAGCGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 1146
QY 541 AATGCGAGCCAAAGCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTCCTTCTCAG 600
Db 1147 AATGCGAGCCAAAGCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTCCTTCTCAG 1206
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTCCAC 660
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTCCAC 1266
QY 661 AGCAGCTACGCGCAGCAGGACGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
Db 1267 AGCAGCTACGCGCAGCAGGACGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 1326
QY 721 CTGTATTACTGACAGAACTCAAAATCAGTCCGGAAGTGCACAAAACAGGACTTGTCTG 780
Db 1327 CTGTATTACTGACAGAACTCAAAATCAGTCCGGAAGTGCACAAAACAGGACTTGTCTG 1386
QY 781 TTTAGCGGTGCTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCTGACCC 840

Db 1387 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCCAGCCAAAACTGGCTACCTGGACCC 1446
Qy 841 TGTATTCCGACAGCGGCTTTCTAAACAAAAAGACAAACAAAGCAATTTTACC 900
Db 1447 TGTATTCCGACAGCGGCTTTCTAAACAAAAAGACAAACAAAGCAATTTTACC 1506
Qy 901 TGGACTGGTGTCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960
Db 1507 TGGACTGGTGTCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 1566
Qy 961 GCTATGGCTCTACAAAAGACGACGAAGAAGTGTCTTTCCATGAGCGGTGTATGATT 1020
Db 1567 GCTATGGCTCTACAAAAGACGACGAAGAAGTGTCTTTCCATGAGCGGTGTATGATT 1626
Qy 1021 TTTGGAAAAGAGCGCCGGAGCTTCAACACATGCAATTTGGACAATGTCTATGATTCAGAC 1080
Db 1627 TTTGGAAAAGAGCGCCGGAGCTTCAACACATGCAATTTGGACAATGTCTATGATTCAGAC 1686
Qy 1081 GAAGAGGAATTTAAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGAGCCGTGGCAGTC 1140
Db 1687 GAAGAGGAATTTAAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGAGCCGTGGCAGTC 1746
Qy 1141 AATTTCCAGACGACGACGACGACCCCTGCGACCGGAGATGTGCATGCTATGGAGCATTA 1200
Db 1747 AATTTCCAGACGACGACGACGACCCCTGCGACCGGAGATGTGCATGCTATGGAGCATTA 1806
Qy 1201 CTTGGCATGGTGGCAAGATAGACGCTGTACCTGCAGGGTCCCATTTGGGCCAAATTT 1260
Db 1807 CTTGGCATGGTGGCAAGATAGACGCTGTACCTGCAGGGTCCCATTTGGGCCAAATTT 1866
Qy 1261 CTTCAACAGATGGACACTTTTCAACCGTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1867 CTTCAACAGATGGACACTTTTCAACCGTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1926
Qy 1321 CGGCTCTCTCAGATCCTCATCAAAAACACGCTGTCTCTGCGAATCTCTCCGGCGGAGTTT 1380
Db 1927 CGGCTCTCTCAGATCCTCATCAAAAACACGCTGTCTCTGCGAATCTCTCCGGCGGAGTTT 1986
Qy 1381 TCAGCTACAAAGTTTGCTTCATTCACCCCAATCTCCACGAGCAAGTGTAGTGGAA 1440
Db 1987 TCAGCTACAAAGTTTGCTTCATTCACCCCAATCTCCACGAGCAAGTGTAGTGGAA 2046
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCTGGAATCCCGAAGTGAGTACACA 1500
Db 2047 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCTGGAATCCCGAAGTGAGTACACA 2106
Qy 1501 TCCAAATTATGCAAAATCTGCCAAGCTTTGATTTTACTGTGGCAACAATGGACTTTTACT 1560
Db 2107 TCCAAATTATGCAAAATCTGCCAAGCTTTGATTTTACTGTGGCAACAATGGACTTTTACT 2166
Qy 1561 GAGCCTCGGCCCATTTGGCACCGGTACCTTACCTTACCGTCCCTGTGA 1604
Db 2167 GAGCCTCGGCCCATTTGGCACCGGTACCTTACCTTACCGTCCCTGTGA 2210

RESULT 5

AD226929

ID AD226929 standard; DNA; 2211 BP.

XX

AC AD226929;

XX

DT 30-JUN-2005 (first entry)

XX

DE Adeno-associated virus DNA SEQ ID NO 79.

XX

KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX Adeno-associated virus.

XX WO200503321-A2.
PN 14-APR-2005.
XX 30-SEP-2004; 2004WO-US028817.
PF 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX (UYPE-) UNIV PENNSYLVANIA.
PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI; 2005-285437/29.
DR New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX Disclosure; SEQ ID NO 79; 569pp; English.
PS The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 & per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;
Query Match 100.0%; Score 1604; DB 14; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 60
Db 607 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 666
Qy 61 AATGCTCAGGAATTTGCATTTGCATTCACATGGCTGGGGCAGAGTCAACACC 120
Db 667 AATGCTCAGGAATTTGCATTTGCATTCACATGGCTGGGGCAGAGTCAACACC 726
Qy 121 AGCACCCGACCTGGGCGCTTGGCCACCTACCAATACCACTTACAAGCAAAATCTCCAGT 180
Db 727 AGCACCCGACCTGGGCGCTTGGCCACCTACCAATACCACTTACAAGCAAAATCTCCAGT 786
Qy 181 GCTTCAACGGGGCGCCAGCAACGACCACTACTTTGGGCTACAGCACCCCTGGGGGTAT 240
Db 787 GCTTCAACGGGGCGCCAGCAACGACCACTACTTTGGGCTACAGCACCCCTGGGGGTAT 846
Qy 241 TTTGATTTCAACAGATTCACATGCGCACTTTTCAACAGCTGAGCTGCGAGGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCACATGCGCACTTTTCAACAGCTGAGCTGCGAGGACTCATCAAC 906
Qy 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCCAAGTCAAG 360
Db 907 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCCAAGTCAAG 966
Qy 361 GAGGTCAACGCAATGATGGCGTCAACACCATCGCTTAATAACCTTTACCAGCACGGTTCAA 420
Db 967 GAGGTCAACGCAATGATGGCGTCAACACCATCGCTTAATAACCTTTACCAGCACGGTTCAA 1026
Qy 421 GTCTTCTCGGACTCGAGTACAGCTTCGGTACGCTTCGGCTCTGGCGCACCGGCTGC 480
Db 1027 GTCTTCTCGGACTCGAGTACAGCTTCGGTACGCTTCGGCTCTGGCGCACCGGCTGC 1086
Qy 481 CTCCCTCGGTTCCCGGCGGACGCTGTTTCATGTTCCGCAATACGCTACTCTGACGCTCAAC 540

Db 1087 |||||CTCCCTCCGTTCCCGCGAGCGTTCATGATTCGCGCAATACGGCTACCTGACGCTCAAC 1146
QY 541 AATGGCAGCAAGCCGCTGGAGCGTTCATCTCTTTACTGCTGGAATATTTCTCTCTCAG 600
Db 1147 AATGGCAGCAAGCCGCTGGAGCGTTCATCTCTTTACTGCTGGAATATTTCTCTCTCAG 1206
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACCTTTTGAAGAACTGCTTTCCAC 660
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACCTTTTGAAGAACTGCTTTCCAC 1266
QY 661 AGCAGCTACGGCAGCAGCCGCTGGACCGGCTGATGAATCTCTCATCAGCAATATAC 720
Db 1267 AGCAGCTACGGCAGCAGCCGCTGGACCGGCTGATGAATCTCTCATCAGCAATATAC 1326
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAAGCACTTGGTG 780
Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAAGCACTTGGTG 1386
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCCAGCCCAAACTGGCTACTGGACCC 840
Db 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCCAGCCCAAACTGGCTACTGGACCC 1446
QY 841 TGTATTTCGAGCAGCGCGTTCATAAAACAAACAGACAAACAAAGCAATTTTACC 900
Db 1447 TGTATTTCGAGCAGCGCGTTCATAAAACAAACAGACAAACAAAGCAATTTTACC 1506
QY 901 TGGACTGGTCTCAAAATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960
Db 1507 TGGACTGGTCTCAAAATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 1566
QY 961 GCTATGGCTTCACACAAAGACGACGAGTGTCTTTCCATGAGCGGTGTGATGATT 1020
Db 1567 GCTATGGCTTCACACAAAGACGAGGAGCAAGTCTTTCCATGAGCGGTGTGATGATT 1626
QY 1021 TTTGAAAAAGAGCCCGGAGCTTCAAAACATGTGCAATTTGGACAAATGATGATTAAGAC 1080
Db 1627 TTTGAAAAAGAGCCCGGAGCTTCAAAACATGTGCAATTTGGACAAATGATGATTAAGAC 1686
QY 1081 GAAGAGGAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
Db 1687 GAAGAGGAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1746
QY 1141 AATTTCCAGACGACGACAGACCCCTGGACCGGAGATGTGATCTATGGGAGCATTA 1200
Db 1747 AATTTCCAGACGACGACAGACCCCTGGACCGGAGATGTGATCTATGGGAGCATTA 1806
QY 1201 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGTCCCAATTTGGGCCAAATTT 1260
Db 1807 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGTCCCAATTTGGGCCAAATTT 1866
QY 1261 CCTCACACAGATGGACATTTTACCCGTCCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1867 CCTCACACAGATGGACATTTTACCCGTCCTCTTATGGCGGCTTTGGACTCAAGAAC 1926
QY 1321 CGCCCTCTCAGATCTCATCAAAAACAGCCCTGTTCTTCGCAATCTCCGGCGGAGTTT 1380
Db 1927 CGCCCTCTCAGATCTCATCAAAAACAGCCCTGTTCTTCGCAATCTCCGGCGGAGTTT 1986
QY 1381 TCAGCTACAAAGTTTGCTTCATTCATCCCAATACTTCCAAGGCAAGTGAAGTGGAA 1440
Db 1987 TCAGCTACAAAGTTTGCTTCATTCATCCCAATACTTCCAAGGCAAGTGAAGTGGAA 2046
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCGCTGGAATCCCAAGTGCAGTACACA 1500
Db 2047 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCGCTGGAATCCCAAGTGCAGTACACA 2106
QY 1501 TCCAAATTATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGGACTTTTACT 1560
Db 2107 TCCAAATTATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGGACTTTTACT 2166
QY 1561 GAGCCTCGCCCAATGGCACCCGTTTACCTTTACCCGCTCCCTGTA 1604
|||||

Db 2167 GAGCCTCGCCCAATGGCACCGTTCACCTTACCCGCTCCCTGTA 2210
RESULT 6
ADW39398
ID ADW39398 standard; DNA; 4347 BP.
XX
AC ADW39398;
XX
DT 24-MAR-2005 (first entry)
XX
DE Adeno-associated virus 2 (AAV2) rep-AAV1 cap fusion DNA - SEQ ID 1.
XX
KW genetically engineered microorganism; vector; rep; cap; ds.
OS Adeno-associated virus 2.
OS Chimeric.
OS Unidentified.
XX
PN CN1461805-A.
XX
PD 17-DEC-2003.
XX
PF 27-MAY-2002; 2002CN-00117965.
XX
PR 27-MAY-2002; 2002CN-00117965.
XX
PA (BENV-) BENYUAN ZHENGYANG GENETIC TECHNOLOGY CO.
XX
PI Wu X, Cao H, Dong X;
XX
DR WPI; 2004-248658/24.
XX
PT Structure of a recombinant herpes simplex virus and its use.
XX
PS Claim 1; SEQ ID NO 1; 102pp; Chinese.
XX
CC The invention relates to a novel recombinant herpes simplex virus (rHSVS) whose genome is respectively inserted by the Adeno-associated virus (AAV) rep gene and cap gene. The AAV genes may be derived from viruses AAV1, AAV2, AAV3, AAV4, AAV5 and AAV6. The methods of the invention may be useful for generating recombinant vectors with high efficiency and no need of reconfiguration. The current sequence is that of the Adeno-associated virus 2 (AAV2) rep-Adeno-associated virus 1 (AAV1) cap fusion DNA (SEQ ID 1) of the invention.
XX
SQ Sequence 4347 BP; 1106 A; 1186 C; 1142 G; 913 T; 0 U; 0 Other;
Query Match 100.0%; Score 1604; DB 13; Length 4347;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGCAATAGAGAGCGCCGACGAGTGGGT 60
Db 2627 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGCAATAGAGAGCGCCGACGAGTGGGT 2686
QY 61 AATGCTCTAGGAAATTTGGCATTCACATGGCTGGCGGACAGAGTTCATCACCACC 120
Db 2687 AATGCTCTAGGAAATTTGGCATTCACATGGCTGGCGGACAGAGTTCATCACCACC 2746
QY 121 AGCACC CGCACCTGGGCTTGGCCACCTACATAACACCTCTTACAAGCAATCTCCAGT 180
Db 2747 AGCACC CGCACCTGGGCTTGGCCACCTACATAACACCTCTTACAAGCAATCTCCAGT 2806
QY 181 GCTTCAACGGGGGCCAGCAACGACCACTACTTGGCTACAGCAACCCCTGGGGGTAT 240
Db 2807 GCTTCAACGGGGGCCAGCAACGACCACTACTTGGCTACAGCAACCCCTGGGGGTAT 2866
QY 241 TTTGATTTTCAACAGATTTCCACTGGCACTTTTCCACACGCTGACTGGCAGGACTCATCAAC 300
Db 2867 TTTGATTTTCAACAGATTTCCACTGGCACTTTTCCACACGCTGACTGGCAGGACTCATCAAC 2926
QY 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGCAAG 360

CC expressing cassette at outer edge of ITR). The DNA sequence is inserted
CC into XbaI site in UL2 or UL44 gene in HSV genome. The DNA sequence of
CC sequence AD246594-AD246598 can also be inserted into other non-essential
CC gene regions in HSV genome. The recombinant HSV is optionally inserted
CC with other DNA sequences homologous with the already-specified fragments.
CC The vectors are for use in gene therapy of e.g. cancer, respiratory
CC diseases and neural diseases. The virus vectors are safe, with long
CC expression time and wide-spectrum of cell infection, even non-cleaved
CC cells and reverse axonal conduction through the incorporated HSV vectors,
CC and high transfer efficiency. The present sequence is the HSV-AAV
CC sequence rep2cap1 for the AAV/HSV vectors of the invention.
XX
SQ Sequence 4347 BP; 1106 A; 1186 C; 1142 G; 913 T; 0 U; 0 Other;
Query Match 100.0%; Score 1604; DB 14; Length 4347;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGGCGTGGCGCACAATGGCAGACAATAACGAAGGGCGCAGAGTGGGT 60
Db 2627 ATGGCTTCAGGGCGTGGCGCACAATGGCAGACAATAACGAAGGGCGCAGAGTGGGT 2686
QY 61 AATGCCCTCAGGAAATTTGGCAATTCGGAATTCACATGCTGGGCGACAGATCATCACC 120
Db 2687 AATGCCCTCAGGAAATTTGGCAATTCGGAATTCACATGCTGGGCGACAGATCATCACC 2746
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACATAACCACTCTACAGCAAAATCTCCAGT 180
Db 2747 AGCACCCGACCTGGGCGCTTGGCCACCTACATAACCACTCTACAGCAAAATCTCCAGT 2806
QY 181 GCTTCAACGGGGGCCAGCAACCAACCACTACTCTGGGCTACAGCAACCCCTGGGGGTAT 240
Db 2807 GCTTCAACGGGGGCCAGCAACCAACCACTACTCTGGGCTACAGCAACCCCTGGGGGTAT 2866
QY 241 TTTGATTTCAACAGATTCACATGCTCCCATCTTTACCAAGTGAATCTGAGGAGTCAATCAAC 300
Db 2867 TTTGATTTCAACAGATTCACATGCTCCCATCTTTACCAAGTGAATCTGAGGAGTCAATCAAC 2926
QY 301 AACAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360
Db 2927 AACAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 2986
QY 361 GAGGTCAACGCAATGATGGCGTCAACCAACCATCGCTTAATAACCTTACAGCAGCGTTCAA 420
Db 2987 GAGGTCAACGCAATGATGGCGTCAACCAACCATCGCTTAATAACCTTACAGCAGCGTTCAA 3046
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGCTAGTCCGCTCTGGGCTCTGGGACACAGGCTGC 480
Db 3047 GTCTTCTCGGACTCGGAGTACAGCTTCGCTAGTCCGCTCTGGGCTCTGGGACACAGGCTGC 3106
QY 481 CTCCCTCCGTTCCGGGGGACGTTTCATGATTCGCAATACGGCTACCTGAGCGTCAAC 540
Db 3107 CTCCCTCCGTTCCGGGGGACGTTTCATGATTCGCAATACGGCTACCTGAGCGTCAAC 3166
QY 541 AATGGCAGCAAGCCGTTGGGACCTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTCAG 600
Db 3167 AATGGCAGCAAGCCGTTGGGACCTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTCAG 3226
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC 660
Db 3227 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC 3286
QY 661 AGCAGTACCGCAGCAGCAGCAGCCTGGACCGGCTGATGATCTCTCATCGACCAATAC 720
Db 3287 AGCAGTACCGCAGCAGCAGCAGCCTGGACCGGCTGATGATCTCTCATCGACCAATAC 3346
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGAGACTTGGCTG 780
Db 3347 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGAGACTTGGCTG 3406
QY 781 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACTGGCTACCTGACCC 840
Db 3407 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACTGGCTACCTGACCC 3466

QY 841 TGTATCGCAGCAGCGCGTCTTCTAAACAAACAAACAGACAAACAGCAATTTTACC 900
Db 3467 TGTATCGCAGCAGCGCGTCTTCTAAACAAACAAACAGACAAACAGCAATTTTACC 3526
QY 901 TGGACTGTGCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960
Db 3527 TGGACTGTGCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 3586
QY 961 GCTATGGCTTCACACAAAGACGACCAAGATTTCTTCCATCAGCGGTGTCTATGATT 1020
Db 3587 GCTATGGCTTCACACAAAGACGACCAAGATTTCTTCCATCAGCGGTGTCTATGATT 3646
QY 1021 TTTGAAAGAGAGCGCGGAGCTTCAACACATGCTGATTTGGACAATGTCTATGATTA 1080
Db 3647 TTTGAAAGAGAGCGCGGAGCTTCAACACATGCTGATTTGGACAATGTCTATGATTA 3706
QY 1081 GAGAGGAAATTAAGGCCACTTAACCTCTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
Db 3707 GAGAGGAAATTAAGGCCACTTAACCTCTGGCCACCGAAGATTTGGGACCGTGGCAGTC 3766
QY 1141 AATTTCCAGAGCAGCAGCAGCAGCCTGCGACCGGAGATGTGATGCTATGGGAGCATTA 1200
Db 3767 AATTTCCAGAGCAGCAGCAGCAGCCTGCGACCGGAGATGTGATGCTATGGGAGCATTA 3826
QY 1201 CTTGCTATGCTGTGGCAAGATAGACAGTGTACCTGCGAGGGTCCCATTTTGGGCCAAAATT 1260
Db 3827 CTTGCTATGCTGTGGCAAGATAGACAGTGTACCTGCGAGGGTCCCATTTTGGGCCAAAATT 3886
QY 1261 CTTCAACAGATGGACACTTTTCCCGGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320
Db 3887 CTTCAACAGATGGACACTTTTCCCGGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 3946
QY 1321 CGGCTCTCTCAGATCTCATCAAAACAGCGCTGTTCTTGGCAATCTCTCCGCGGAGTTT 1380
Db 3947 CGGCTCTCTCAGATCTCATCAAAACAGCGCTGTTCTTGGCAATCTCTCCGCGGAGTTT 4006
QY 1381 TCAGCTCAAAAGTTTGTCTTCAATTCATCCCAATACTCCAGGCAAGTGTGTGGAA 1440
Db 4007 TCAGCTCAAAAGTTTGTCTTCAATTCATCCCAATACTCCAGGCAAGTGTGTGGAA 4066
QY 1441 ATTGAATGGAGCTCGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 4067 ATTGAATGGAGCTCGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 4126
QY 1501 TCCAAATTATGCAAAATCTGCCAAAGTGTGATTTTACTGTGCAACAATGGACTTTTACT 1560
Db 4127 TCCAAATTATGCAAAATCTGCCAAAGTGTGATTTTACTGTGCAACAATGGACTTTTACT 4186
QY 1561 GAGCCTCGCCCAATTTGGCACCCTTACCTTACCCTCCCTGTGA 1604
Db 4187 GAGCCTCGCCCAATTTGGCACCCTTACCTTACCCTCCCTGTGA 4230

RESULT 8

AAD00772

ID AAD00772 standard; DNA; 4718 BP.

AC AAD00772;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 DNA.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; ss.

XX Adeno associated virus serotype 1.

XX Key Location/Qualifiers

FT repeat_unit 1..143

FT /*tag= a

```

FT /label= 5' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT 89..110
FT /tag= b
FT /bound moiety= "Rep protein"
FT 124..125
FT /tag= c
FT /note= "Terminal resolute site (TRS)"
FT 219..226
FT /tag= d
FT /bound moiety= "USF"
FT 236..299
FT /tag= e
FT /label= P5_promoter
FT 237..245
FT /tag= f
FT /bound moiety= "YY1 factor"
FT 270..275
FT /tag= g
FT /label= P5_TATA-Box
FT 299..306
FT /tag= h
FT /note= "YY1/p5 RNA"
FT 335..2272
FT /tag= j
FT /product= "Rep 68"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT 335..2206
FT /tag= i
FT /product= "Rep 78"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT 857..862
FT /tag= l
FT /label= P19_TATA_Box
FT 882..883
FT /tag= m
FT /note= "P19 RNA"
FT 1007..2272
FT /tag= o
FT /product= "Rep 40"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT 1007..2206
FT /tag= n
FT /product= "Rep 52"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT 1836..1841
FT /tag= p
FT /label= P40_TATA-BOX
FT 1875..1876
FT /tag= q
FT /note= "P40 RNA"
FT 1924..2220
FT /tag= k
FT /note= "This region interrupts the coding sequence of Rep
FT 68 and Rep 40"
FT 2223..4433
FT /tag= r
FT /product= "VP1 protein"
FT /note= "Capsid protein"
FT 2634..4433
FT /tag= s
FT /product= "VP2 protein"
FT /note= "Capsid protein"
FT /partial
FT 2829..4433

```

```

FT /tag= t
FT /product= "VP3 protein"
FT /note= "Capsid protein"
FT 4447..4452
FT /tag= u
FT repeat_unit
FT 4576..4718
FT /tag= v
FT /label= 3' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT
XX WO200028061-A2.
XX 18-MAY-2000.
XX
XX 02-NOV-1999; 99WO-US025694.
XX
XX 05-NOV-1998; 98US-0107114P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Xiao W;
XX
XX WPI; 2000-376571/32.
XX P-RSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
XX AAY71169.
XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
XX preparation of medicament for delivery of a transgene to a host.
XX
XX Claim 1; Fig 1; 108pp; English.
XX
XX The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
XX characterised by two inverted terminal repeats (ITR) and open reading
XX frames for rep and capsid (cap) proteins. The rep reading frame encodes
XX four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
XX frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
XX sequence or its fragments particularly ITRs, rep and cap coding regions,
XX are useful in production of recombinant viral vectors for gene delivery.
XX These vectors can be used as gene therapy vectors, vaccine vectors or
XX antisense delivery vectors. The AAV-1 does not induce the formation of
XX neutralising antibodies specific to any serotype of AAV hence is useful
XX for transforming host cells, and in the preparation of a medicament for
XX the delivery of transgene to a host
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 1604; DB 3; Length 4718;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ATGGCTTCAGCGGTGGCGCCACCAATGGCAGACAATAACGAAGGCGCGACGAGTGGGT 60
Db 2829 ATGGCTTCAGCGGTGGCGCCACCAATGGCAGACAATAACGAAGGCGCGACGAGTGGGT 2888
QY 51 AATGCCTCAGGAATGGCATTCGATTCACATGGCTGGGCGCAGAGTCAATCCACC 120
Db 2889 AATGCCTCAGGAATGGCATTCGATTCACATGGCTGGGCGCAGAGTCAATCCACC 2948
QY 121 AGCACCCGCACCTGGGCGCTTGCCACCTACAAATAACCACTCTACAAGCAAAATCTCCAGT 180
Db 2949 AGCACCCGCACCTGGGCGCTTGCCACCTACAAATAACCACTCTACAAGCAAAATCTCCAGT 3008
QY 181 GCTTCAACGGGGGCCAGCAACGACAAACCACTACTTCGGGTACAGCACCCCTCGGGGGTAT 240
Db 3009 GCTTCAACGGGGGCCAGCAACGACAAACCACTACTTCGGGTACAGCACCCCTCGGGGGTAT 3068
QY 241 TTTGATTTCAACAGATTCCACCTGCGACCTTTTCCACCGTACTGCGAGCTCATCAAC 300
Db 3069 TTTGATTTCAACAGATTCCACCTGCGACCTTTTCCACCGTACTGCGAGCTCATCAAC 3128
QY 301 AACAAATTGGGGATTCCGGCCCAAGAGACTCAACTTCAAACCTCTTCAAACATCAAGTCAAG 360

```

Db 3129 AACAAATGGGGATTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 3188
QY 361 GAGGTCAACGACGAATGATGGCGTCACAAACCATCGCTAAATAACCTTACCACGACGGTTCAA 420
Db 3189 GAGGTCAACGACGAATGATGGCGTCACAAACCATCGCTAAATAACCTTACCACGACGGTTCAA 3248
QY 421 GTCTTCTCGGACTCGGAGTACCAAGCTTCGGTAGAGTCTCGGTCTCGGTCTCGGCACCAAGGCTGC 480
Db 3249 GTCTTCTCGGACTCGGAGTACCAAGCTTCGGTAGAGTCTCGGTCTCGGTCTCGGCACCAAGGCTGC 3308
QY 481 CTCCTCTCGGTTCCTCGGCGGACGGTTCATGATTCGCGAATATTCGGCTACGCTGACGCTCAAC 540
Db 3309 CTCCTCTCGGTTCCTCGGCGGACGGTTCATGATTCGCGAATATTCGGCTACGCTGACGCTCAAC 3368
QY 541 AATGGCAGCAACCCCGTGGGACCTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 600
Db 3369 AATGGCAGCAACCCCGTGGGACCTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 3428
QY 601 ATGCTGAGAACCGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 660
Db 3429 ATGCTGAGAACCGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 3488
QY 661 AGCAGCTACGCGACGACGACGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 3489 AGCAGCTACGCGACGACGACGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 3548
QY 721 CTGTATTACTCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGAAACCAAGGACTTGGCTG 780
Db 3549 CTGTATTACTCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGAAACCAAGGACTTGGCTG 3608
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACTCGGTACTCGTGAACCC 840
Db 3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACTCGGTACTCGTGAACCC 3668
QY 841 TGTATTCCGACAGCGCGCTTCTTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 900
Db 3669 TGTATTCCGACAGCGCGCTTCTTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 3728
QY 901 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
Db 3729 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 3788
QY 961 GCTATGGCTCTCACAAAGACGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 1020
Db 3789 GCTATGGCTCTCACAAAGACGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 3848
QY 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATTACAGAC 1080
Db 3849 TTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATTACAGAC 3908
QY 1081 GAAGAGGAATTAAGCCACTAACCTCTGGCCACCGAAAGATTTGGACCGTGGCAGTC 1140
Db 3909 GAAGAGGAATTAAGCCACTAACCTCTGGCCACCGAAAGATTTGGACCGTGGCAGTC 3968
QY 1141 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCATGCTATGGGAGCATTA 1200
Db 3969 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCATGCTATGGGAGCATTA 4028
QY 1201 CTTGGCATGTGTGGCAAGATAGAGAGCTGTACTGAGAGGTGCCATTTGGGCCAAATTT 1260
Db 4029 CTTGGCATGTGTGGCAAGATAGAGAGCTGTACTGAGAGGTGCCATTTGGGCCAAATTT 4088
QY 1261 CCTCACACATGACACTTTTCAACCGTCTCTTATATGGCGGCTTTGGACATCAAGAAC 1320
Db 4089 CCTCACACATGACACTTTTCAACCGTCTCTTATATGGCGGCTTTGGACATCAAGAAC 4148
QY 1321 CCGCTCTCAGATCTCTCATCAAAAAACACGCTCTTCTCGAATCTCTCGGCGGAGTTT 1380
Db 4149 CCGCTCTCAGATCTCTCATCAAAAAACACGCTCTTCTCGAATCTCTCGGCGGAGTTT 4208
QY 1381 TCAGCTACAAAGTTTGCTTCATTCATCACCCCAATCTCCACAGACAAGTGAAGTGGAA 1440
Db 4209 TCAGCTACAAAGTTTGCTTCATTCATTCACCCCAATCTCCACAGACAAGTGAAGTGGAA 4268

QY 1441 ATTGAATGGGAGCTCGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 4269 ATTGAATGGGAGCTCGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 4328
QY 1501 TCCAAATTATGCAAAATCTGCCAACCTTGATTTTACTGTGGACAAACAATGGACTTTTATCT 1560
Db 4329 TCCAAATTATGCAAAATCTGCCAACCTTGATTTTACTGTGGACAAACAATGGACTTTTATCT 4388
QY 1561 GAGCCTCGCCCCATTGGCACCCCGTTACCTTTACCCGTCCCTCTGTA 1604
Db 4389 GAGCCTCGCCCCATTGGCACCCCGTTACCTTTACCCGTCCCTCTGTA 4432
RESULT 9
ADE76507
ID ADE76507 standard; DNA; 4718 BP.
XX ADE76507;
AC ADE76507;
XX
DT 29-JAN-2004 (first entry)
XX
DE Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
XX
DE adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW dermatological; antiinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
OS Adeno-associated virus 1.
XX
PN EP1310571-A2.
XX
PD 14-MAY-2003.
XX
PF 12-NOV-2002; 2002BP-00257826.
XX
PR 13-NOV-2001; 2001US-0350607P.
PR 17-DEC-2001; 2001US-0341117P.
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386675P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Gao G, Wilson JM, Alvira M;
XX
XX WPI; 2003-450984/43.
XX
PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
PS Claim 14; SEQ ID NO 6; 419pp; English.
XX
CC The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antirheumatic, antiarthritic, neuroprotective,
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;

Query Match 100.0%; Score 1604; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGCGAGTGGGT 60
DB 2829 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGCGAGTGGGT 2888

QY 61 AATGCTTCAGGAAATTTGGCAATTTGCGATTTCCATGCTGGGGGACAGAGTTCATCAACCACC 120
DB 2889 AATGCTTCAGGAAATTTGGCAATTTGCGATTTCCATGCTGGGGGACAGAGTTCATCAACCACC 2948

QY 121 AGCACCGCACCTGGGCTTGGCCACCTTACATAAACCACCTCTACAGCAAAATCTCCAGT 180
DB 2949 AGCACCGCACCTGGGCTTGGCCACCTTACATAAACCACCTCTACAGCAAAATCTCCAGT 3008

QY 181 GCTTCAACGGGGCCAGCAACACACCACTACTTTCGGCTACAGACCCCTGGGGGTAT 240
DB 3009 GCTTCAACGGGGCCAGCAACACCACTACTTTCGGCTACAGACCCCTGGGGGTAT 3068

QY 241 TTTGATTTCAACAGATTTCCATGCTGCACTTTTCCACCACTGACTGGCAGCACTCATCAAC 300
DB 3069 TTTGATTTCAACAGATTTCCATGCTGCACTTTTCCACCACTGACTGGCAGCACTCATCAAC 3128

QY 301 AACAAATTTGGGAATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAAGTCAAG 360
DB 3129 AACAAATTTGGGAATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAAGTCAAG 3188

QY 361 GAGGTCAACAGCAATGATGGCGTCAACACCATCGCTTAATACTTACAGACGGTTCAA 420
DB 3189 GAGGTCAACAGCAATGATGGCGTCAACACCATCGCTTAATACTTACAGACGGTTCAA 3248

QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTTCGCGACCAAGGCTGC 480
DB 3249 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTTCGCGACCAAGGCTGC 3308

QY 481 CTCCCTCCGCTTCGGCGGACGCTGTTTCATGATTCGGCAATACGGCTACCTGAGCTCAAC 540
DB 3309 CTCCCTCCGCTTCGGCGGACGCTGTTTCATGATTCGGCAATACGGCTACCTGAGCTCAAC 3368

QY 541 AATGGCAGCAAGCGGTGGGAGCTTTCATCTTTTACTGCTGGAATATTTTCCCTTCTCAG 600
DB 3369 AATGGCAGCAAGCGGTGGGAGCTTTCATCTTTTACTGCTGGAATATTTTCCCTTCTCAG 3428

QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTTCCAC 660
DB 3429 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTTCCAC 3488

QY 661 AGCAGCTACGGCACAGCCAGAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
DB 3489 AGCAGCTACGGCACAGCCAGAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 3548

QY 721 CTGTATTACTGAACAGAACTCAAATCAGTCCGGAAGTCCCAAAACAGAGACTTGTCTG 780
DB 3549 CTGTATTACTGAACAGAACTCAAATCAGTCCGGAAGTCCCAAAACAGAGACTTGTCTG 3608

QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACCTGGACCC 840
DB 3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACCTGGACCC 3668

QY 841 TGTATTCCGCGAGCGCGCTTTCTAATAACAAAACAGACAAACAAAGCAATTTTACC 900
DB 3669 TGTATTCCGCGAGCGCGCTTTCTAATAACAAAACAGACAAACAAAGCAATTTTACC 3728

QY 901 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960
DB 3729 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 3788

QY 961 GCTATGGCTCAACAAAAGACGAGAGACAAGTCTTTTCCATGAGCGGTGTCATGATT 1020
DB 3789 GCTATGGCTCAACAAAAGACGAGAGACAAGTCTTTTCCATGAGCGGTGTCATGATT 3848

QY 1021 TTTGGAAAGAGAGCGCGGAGCTTCAAAACACTGCATTTGGCAATGTCTATGATTACAGAC 1080
DB 3849 TTTGGAAAGAGAGCGCGGAGCTTCAAAACACTGCATTTGGCAATGTCTATGATTACAGAC 3908

QY 1081 GAAGAGGAAATTAAGAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
DB 3909 GAAGAGGAAATTAAGAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 3968

QY 1141 AATTTCAGACAGCAGACAGACCTTGCACCGGAGATGTGCATGTCTATGGAGCATTA 1200
DB 3969 AATTTCAGACAGCAGACAGACCTTGCACCGGAGATGTGCATGTCTATGGAGCATTA 4028

QY 1201 CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTGGGCCCAAAATT 1260
DB 4029 CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTGGGCCCAAAATT 4088

QY 1261 CCTCACACAGATGACACTTTTCAACCGGTCTCTCTTATGGCGGCTTTGGGACTCAAGAAC 1320
DB 4089 CCTCACACAGATGACACTTTTCAACCGGTCTCTCTTATGGCGGCTTTGGGACTCAAGAAC 4148

QY 1321 CGGCTCTCAGATCTCTCATCAAAACACGCTGTCTCTGCGAATCTCTCCGGCGAGTTT 1380
DB 4149 CGGCTCTCAGATCTCTCATCAAAACACGCTGTCTCTGCGAATCTCTCCGGCGAGTTT 4208

QY 1381 TCAGCTACAAAGTTTGTCTTCAATTCATCACCACTCACTCCAGGACCAAGTGCAGTACACA 1440
DB 4209 TCAGCTACAAAGTTTGTCTTCAATTCATCACCACTCACTCCAGGACCAAGTGCAGTACACA 4268

QY 1441 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
DB 4269 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 4328

QY 1501 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAACAATGGGACTTTATACT 1560
DB 4329 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAACAATGGGACTTTATACT 4388

QY 1561 GAGCTCGCCCATTTGGCACCCGTTACCTTACCGTCCCTGTA 1604
DB 4389 GAGCTCGCCCATTTGGCACCCGTTACCTTACCGTCCCTGTA 4432

RESULT 10
ADL13984
ID ADL13984 standard; DNA; 4718 BP.
XX
AC ADL13984;
XX
DT 06-MAY-2004 (first entry)
XX
DE Adeno-associated virus serotype 1 complete DNA.
XX
KW ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;
KW expression construct; adeno-associated virus;
KW integration efficiency element; inverted terminal repeat; integration;
KW chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;
KW retinoblastoma; inflammatory disease; arthritis;
KW neurodegenerative disease.
XX
OS Adeno-associated virus 1.
XX
PN WO2003087334-A2.
XX
PD 23-OCT-2003.
XX
PF 09-APR-2003; 2003WO-US011191.
XX
PR 09-APR-2002; 2002US-0371044P.
XX
PA (CORR) CORNELL RES FOUND INC.
XX
PI Falck-Pedersen ES, Philipott N;
XX
DR WPI; 2003-833723/77.

XX New expression construct comprising a nucleic acid sequence encoding an
PT adeno-associated virus integration efficiency element, useful for
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease
PT e.g. arthritis.
XX
XX Disclosure; SEQ ID NO 3; 62pp; English.
XX
CC The invention relates to an expression construct comprising a nucleic
CC acid sequence encoding an adeno-associated virus integration efficiency
CC element (AAV IRE), which is devoid of AAV inverted terminal repeats (AAV
CC ITRs) and site-specifically integrates into a host cell chromosome when
CC provided to the host cell in conjunction with an AAV Rep protein. The
CC expression construct can be used as a therapeutic factor for treating a
CC mammal for a pathologic state which is cancer, including lung cancer,
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,
CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic
CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic
CC state includes inflammatory disease (arthritis), neurodegenerative
CC disease, a disease of an organ attributed to the presence of increased or
CC decreased level of a particular gene product(s). This sequence
CC corresponds to the AAV serotype 1 complete DNA including the IEE
CC sequence.
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 1604; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGCGTGGCGCCACCAATGGCAGACAATAACGAAGCGCCGACGAGTGGGT 60
DB 2829 ATGGCTTCAGGCGTGGCGCCACCAATGGCAGACAATAACGAAGCGCCGACGAGTGGGT 2888
QY 61 AATGCTCAGGAAATGGCAATGCGATTCCACATGGCTGGCGCAGAGTCAATCACCAC 120
DB 2889 AATGCTCAGGAAATGGCAATGCGATTCCACATGGCTGGCGCAGAGTCAATCACCAC 2948
QY 121 AGCACCCGACCTGGGCTTGGCCACCTCAATPAACCACTCTTCAAGCAAAATCTCCAGT 180
DB 2949 AGCACCCGACCTGGGCTTGGCCACCTCAATPAACCACTCTTCAAGCAAAATCTCCAGT 3008
QY 181 GCTTCAACGGGGCCAGCAACCACTACTTCCGGCTACAGACCCCTGGGGGTAT 240
DB 3009 GCTTCAACGGGGCCAGCAACCACTACTTCCGGCTACAGACCCCTGGGGGTAT 3068
QY 241 TTTGATTTCAACAGATTCACATGCCACTTTTACCACGCTGACTGGCAGGAGCTCATCAAC 300
DB 3069 TTTGATTTCAACAGATTCACATGCCACTTTTACCACGCTGACTGGCAGGAGCTCATCAAC 3128
QY 301 AACAAATGGGGATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360
DB 3129 AACAAATGGGGATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 3188
QY 361 GAGGTCAACGAAATGATGGCGTCAACCAATCGCTAATAAATCTTACCGAGCAGCGTTCAA 420
DB 3189 GAGGTCAACGAAATGATGGCGTCAACCAATCGCTAATAAATCTTACCGAGCAGCGTTCAA 3248
QY 421 GTCTCTCGGACTCGGAGTACAGCTTCCGCAATACCGGCTACCTGAGCGCTCAAC 540
DB 3249 GTCTCTCGGACTCGGAGTACAGCTTCCGCAATACCGGCTACCTGAGCGCTCAAC 3308
QY 481 CTCCTCTCCGTTCCCGCGGAGCTGTTTCATGATTTCCGCAATACCGGCTACCTGAGCGCTCAAC 540
DB 3309 CTCCTCTCCGTTCCCGCGGAGCTGTTTCATGATTTCCGCAATACCGGCTACCTGAGCGCTCAAC 3368
QY 541 AATGGCAGCAACCGGTGGAGCTTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 600
DB 3369 AATGGCAGCAACCGGTGGAGCTTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 3428

RESULT 11
ADG39758
ID ADG39758 standard; DNA; 4718 BP.

QY 601 ATGCTGAGAACGGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660
DB 3429 ATGCTGAGAACGGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 3488
QY 661 AGCAGCTACGGCGCAGCAGCGCTGGACCGGCTGATGAATCCTCTCATCGCAAAATAC 720
DB 3489 AGCAGCTACGGCGCAGCAGCGCTGGACCGGCTGATGAATCCTCTCATCGCAAAATAC 3548
QY 721 CTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAGAGACTTGTCTG 780
DB 3549 CTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAGAGACTTGTCTG 3608
QY 781 TTTAGCCGTGGTCTCCAGCTGGCGATGCTGTTCAGCCCAAAACTGGCTACTCGGACCC 840
DB 3609 TTTAGCCGTGGTCTCCAGCTGGCGATGCTGTTCAGCCCAAAACTGGCTACTCGGACCC 3668
QY 841 TGTTATCCGCGACGCGGCTTTCTAAACAAAAACAGAAACAGAACTCAATCAACAGCAATTTTACC 900
DB 3669 TGTTATCCGCGACGCGGCTTTCTAAACAAAAACAGAAACAGAACTCAATCAACAGCAATTTTACC 3728
QY 901 TGGACTGTGTCTTCAAAATATAAATCAATGGGCGTGAATCCATCATCAACCCCTGGCACT 960
DB 3729 TGGACTGTGTCTTCAAAATATAAATCAATGGGCGTGAATCCATCATCAACCCCTGGCACT 3788
QY 961 GCTATGGCTCTACACAAAGACGACGAAGTCTTTTCCCATGAGCGGTGTCTGATTT 1020
DB 3789 GCTATGGCTCTACACAAAGACGACGAAGTCTTTTCCCATGAGCGGTGTCTGATTT 3848
QY 1021 TTTGAAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGGCAAGTCTTGGGCAAGTCTTGGGCAAGTCT 1080
DB 3849 TTTGAAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGGCAAGTCTTGGGCAAGTCTTGGGCAAGTCT 3908
QY 1081 GAAGAGGAAATTAAGAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGCAAGTCTTGGGCAAGTCT 1140
DB 3909 GAAGAGGAAATTAAGAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGCAAGTCTTGGGCAAGTCT 3968
QY 1141 AATTTCCAGAGCAGCAGCAGCAGCCTTGCACCGAGAGTGTGCATGTCTATGGGAGCATTTA 1200
DB 3969 AATTTCCAGAGCAGCAGCAGCAGCCTTGCACCGAGAGTGTGCATGTCTATGGGAGCATTTA 4028
QY 1201 CCTGCGATGGTGTGCGAAGATAGACGCTGATACCTGCGAGGTCCTCAATTTGGGCAAGTCTTGGGCAAGTCT 1260
DB 4029 CCTGCGATGGTGTGCGAAGATAGACGCTGATACCTGCGAGGTCCTCAATTTGGGCAAGTCTTGGGCAAGTCT 4088
QY 1261 CCTCACAAGATGGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTTGGGCAAGTCTTGGGCAAGTCT 1320
DB 4089 CCTCACAAGATGGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTTGGGCAAGTCTTGGGCAAGTCT 4148
QY 1321 CCGCCTCTCTCAGATCCTCATCAAAACAGCGCTGTTCCTGCGAATCTCTCCGCGGAGTTT 1380
DB 4149 CCGCCTCTCTCAGATCCTCATCAAAACAGCGCTGTTCCTGCGAATCTCTCCGCGGAGTTT 4208
QY 1381 TCAGCTACAAAGTTGCTTTCATCACCCTGATCTCAGAGGACAGTGTGAGTGGAA 1440
DB 4209 TCAGCTACAAAGTTGCTTTCATCACCCTGATCTCAGAGGACAGTGTGAGTGGAA 4268
QY 1441 ATTGAATGGGAGCTGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
DB 4269 ATTGAATGGGAGCTGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 4328
QY 1501 TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGCAACAAATGGACTTTTATCT 1560
DB 4329 TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGCAACAAATGGACTTTTATCT 4388
QY 1561 GAGCCTCGCCCCATTTGGGACCGCTTACCTTACCCGCTCCCTCTGTA 1604
DB 4389 GAGCCTCGCCCCATTTGGGACCGCTTACCTTACCCGCTCCCTCTGTA 4432

1201	Qy	CCTGGCATGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCACTTTGGGCGCAAAATT	1260
4029	Db	CCTGGCATGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCACTTTGGGCGCAAAATT	4088
1261	Qy	CCTCACACAGATGGACACTTTTCAACCGCTCTCTTTATGGCGCGCTTTTGACATCAAGAAC	1320
4089	Db	CCTCACACAGATGGACACTTTTCAACCGCTCTCTTTATGGCGCGCTTTTGACATCAAGNAC	4148
1321	Qy	CCGCTCTCTCAGATCTCATCAAAAAACACGCTGTCTCTCGCAATCTCTCGGCGGAGTTT	1380
4149	Db	CCGCTCTCTCAGATCTCATCAAAAAACACGCTGTCTCTCGCAATCTCTCGGCGGAGTTT	4208
1381	Qy	TCAGCTACAAAGTTTGCTTCATTCACCCCAATCTCCAACAGACAAGTGAGTGTGGAA	1440
4209	Db	TCAGCTACAAAGTTTGCTTCATTCATCAACCAATCTCCAACAGACAAGTGAGTGTGGAA	4268
1441	Qy	ATTGMAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA	1500
4269	Db	ATTGMAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA	4328
1501	Qy	TCCAAATTATGCAAAATCTGCCAACGTTTGATTTTACTGTGGACAAACAATGGACTTTATACT	1560
4329	Db	TCCAAATTATGCAAAATCTGCCAACGTTTGATTTTACTGTGGACAAACAATGGACTTTATACT	4388
1561	Qy	GAGCCTCGCCCCAATGGCAACCGGTACTCTTACCGTCCCTCTGTA	1604
4389	Db	GAGCCTCGCCCCAATGGCAACCGGTACTCTTACCGTCCCTCTGTA	4432

RESULT 12

ACC58477

ACC58477
ID ACC58477 standard: DNA: 7447 BP.

AC ACC58477:

26-AUG-2003 (first entry)

DE Vector pFBDAAV1-VPP with adeno associated virus 2 VP capsid gene.

AAV: pFBDAAV1-VPP: capsid: insect: gene therapy: vector: ss:

XX Adeno associated virus serotype 1.

OS **synthetic.**

PN WO2003042361-A2.

22-MAY-2003

XX
PF 08-NOV-2002: 2002WO-US035829.XX
PB 09-NOV-2001: 2001US-00986618

03-NOV-2001; 2001US-00388918:
13-AUG-2002; 2002US-00216870:
PR

XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.XX
PT Korin RM. Urabe M. Ding C:

WPB: 2003-449568/42

xx Producing an adeno-associated virus (AAV) in an insect cell, e.g. for
 pt gene therapy, comprises introducing an insect cell-compatible vector into
 pt an insect cell and maintaining the insect cell under conditions that
 pt produce AAV.

Example 7: Page 68-72: 84pp: English.

CC The present sequence is that of vector pFBDAAV1-VP, which comprises a
CC modified adeno associated virus serotype 1 (AAV1) VP capsid gene. The
CC modified VP gene has the VP1 initiation codon mutated to ACG to reduce
CC its translation efficiency. An out-of-frame ATG, which diminishes the
CC translation of VP2 and VP3, is also modified, and the splice acceptor
CC site downstream of the VP1 initiation codon has been destroyed.
CC Recombinant baculovirus vectors comprising the mutant AAV1 VP gene were

produced from pFBDAAV1-VP for use in the production of recombinant AAV particles comprising AAV1 capsid proteins in insect (Sf9) cells. The method of producing AAV in insect cells provides an efficient, safe and economical means of producing a large amount of recombinant AAV particles which may be used in gene therapy

Sequence 7447 BP; 1870 A; 1934 C; 1795 G; 1848 T; 0 U; 0 Other;

Query Match 100.0%; Score 1604; DB 8; Length 7447;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTTCAGCGGTGGCGACCAATGGGAGACAATAACGAAGGCCCGACGAGTGGGT 60

Dh 5227 ATGGCTTCAGCGGTGGCGACCAATGGGAGACAATAACGAAGGCCCGACGAGTGGGT 5286

QY 61 AATGCCTCAGGAAATTTGGCATTTCGATTTCCACATGGCTGGGCGACAGAGTCATCACACC 120

QY 121 AGCACCGCACCTGGGCCTTGGCCACCTACAATAACCACTCTACAGCAAAATCTCCAGT 180

181 GTTCAACGGGGGCCAGCAACCACTACTTCGGCTACAGCACCCCTCGGGGGTAT 240

DB	5407	241	300
QY	GCTTCAACGGGGGGCCAGCAACGACACACATACATTCGGCTACAGACACACCTCTCTGGGGGGAT	TTTGATTTCAACAGATTCCACTGCCCACTTTTCCACCACTGACTGGCAGCGACTCATCAAC	

QY	301	AA	CAAT	GGG	GANT	TCGG	CCCA	GAG	ACT	AA	CTT	CAAA	CTT	TTC	AAC	ATC	CAAG	TC	AA	GT	CAAG	360	
DB	5467	TTT	GAT	TTT	CA	CAG	ATT	CC	CA	CT	CTT	TTC	CC	CA	CC	CG	CT	TTT	TTC	CA	CC	CA	552

Db	5527	AACAAATGGGGGAAATTCGGCCCAAGAGACTCAACTTCAACATCTTCAACATCCAAAGTCAAG
Ov	361	GAGCTCACGACGAATGATGGCGTCAACAACCATCGCTTAATAAAGTTTACCAGCAGGGTTCAA

Db 5587 GAGGTACAGCAATGATGGCGTCACAAACCATCGCTTAATAACCTTACGACGACGGTTCAA 5646

Db 5647 GTCTTCGGACTCGGAGTACCAAGCTTCGGTACGTCCTCGGTCTCGGCACAGGGGTGC 5706

5707 CTCCTCCGTTCCCGCGACGTGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 5766

QY	541	5767	5826
ATGGGAGCCCAAGCCCGTGGGAGCTTCAATCCCTTTACATGCCCTGGAAATATTTCCCTCTCTCAG	ATGGGAGCCCAAGCCCGTGGGAGCTTCAATCCCTTTACATGCCCTGGAAATATTTCCCTCTCTCAG	ATGGGAGCCCAAGCCCGTGGGAGCTTCAATCCCTTTACATGCCCTGGAAATATTTCCCTCTCTCAG	

Qy	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000
Db	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	601	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	602	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	603	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	604	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	605	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	606	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	607	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCCAC	608	ATGCTGAGAACGGGCAACAACTTACCTT	CAGCTACACCTTTGAGGAAGTGCTTTCC																																																																																																																																																																																																																																																																																																																																																																																						

Qy	661	ACGAGCTACGCGCACAGCCAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC	720
Dh	5887	AGCAGTTTACGCGCACAGCCAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC	5946

QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGAGTGCCCAAAACAAGACTTGTGTG 780

Qy 781 TTTAGCGTGGGTCTCCAGCTGGCATGCTCTGTTCAGCCCAAACTGGCTACTCTGGACC 840

841 TGTATTGGCAGCAGCGGTTCTTAAAAA CAAAAACAGACAACAA CAGCAATTTTACC 900

Db 6127 TGGAGTGGTCTTCAAAATATTAACCTCAATGGCGGTGAATCATCATCAACCTGGCACT 6186
Qy 961 GCTATGGCTCACAACAGACGAGCAAGTCTTTCCCATGAGCGGTGTCTATGATT 1020
Db 6187 GCTATGGCTCACAACAGACGAGCAAGTCTTTCCCATGAGCGGTGTCTATGATT 6246
Qy 1021 TTTGGAAAAGAGCGCCGAGCTTCAAACTGCAATTTGGACAATGTCTATGATTACAGAC 1080
Db 6247 TTTGGAAAAGAGCGCCGAGCTTCAAACTGCAATTTGGACAATGTCTATGATTACAGAC 6306
Qy 1081 GAAGAGGAATTAAGCCACTTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
Db 6307 GAAGAGGAATTAAGCCACTTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 6366
Qy 1141 AATTTCCAGACGACGACAGACACCTCGGACCGGAGATGTGCTATGAGGACATT 1200
Db 6367 AATTTCCAGACGACGACAGACACCTCGGACCGGAGATGTGCTATGAGGACATT 6426
Qy 1201 CTTGGCATGTTGGCAAGATAGAGACGTTGCTGACGGGTCCCATTTGGGCCAAAATT 1260
Db 6427 CTTGGCATGTTGGCAAGATAGAGACGTTGCTGACGGGTCCCATTTGGGCCAAAATT 6486
Qy 1261 CTTCAACAGATGGACACTTTTACCCGCTCTCTCTTTATGGGGGCTTTGGACTCAAGAAC 1320
Db 6487 CTTCAACAGATGGACACTTTTACCCGCTCTCTCTTTATGGGGGCTTTGGACTCAAGAAC 6546
Qy 1321 CGCCCTCTCAGATCTCATCAAAAACACGCTGTTTCTCGGAATCTCTCGGGGGAGTTT 1380
Db 6547 CGCCCTCTCAGATCTCATCAAAAACACGCTGTTTCTCGGAATCTCTCGGGGGAGTTT 6606
Qy 1381 TCAGCTACAAGTTTCTCTATCTATCAACCAATCTCCAGGACAAAGTGTGGA 1440
Db 6607 TCAGCTACAAGTTTCTCTATCTATCAACCAATCTCCAGGACAAAGTGTGGA 6666
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 6667 ATTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAGTACACA 6726
Qy 1501 TCCAAATTATGCAAAATCTGCCAAGTGTGATTTTACTGTGGACAAATGGACTTTATCT 1560
Db 6727 TCCAAATTATGCAAAATCTGCCAAGTGTGATTTTACTGTGGACAAATGGACTTTATCT 6786
Qy 1561 GAGCTCGCCCAATTGGACCCGTTACCTTACCCGTCCTGTA 1604
Db 6787 GAGCTCGCCCAATTGGACCCGTTACCTTACCCGTCCTGTA 6830

RESULT 13

AD226931 standard; DNA; 2211 BP.

XX AC AD226931;

XX DT 30-JUN-2005 (first entry)

XX DE Adeno-associated virus DNA SEQ ID NO 81.

XX KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
XX KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
XX KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
XX KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
XX KW antipeptidic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

XX OS Adeno-associated virus.

XX PN W020503321-A2.

XX PD 14-APR-2005.

XX PF 30-SEP-2004; 2004MO-US028817.

XX

PR 30-SEP-2003; 2003US-0508236P.
PR 29-APR-2004; 2004US-0566546P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI; 2005-285437/29.
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
PS Claim 19; SEQ ID NO 81; 569pp; English.
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
SQ Sequence 2211 BP; 582 A; 653 C; 535 G; 441 T; 0 U; 0 Other;

Query Match 99.8%; Score 1600.8; DB 14; Length 2211;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1602; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 60
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 666
Qy 61 AATGCTCAGGAATTTGGCATTTGGATTCACATGGCTGGGGGACAGAGTATCAACACC 120
Db 667 AATGCTCAGGAATTTGGCATTTGGATTCACATGGCTGGGGGACAGAGTATCAACACC 726
Qy 121 AGCACCCGACCTGGGCTTGGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT 180
Db 727 AGCACCCGACCTGGGCTTGGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT 786
Qy 181 GCTTCAACGGGGGCGCAGCAACCACTACTTTGGCTACAGCACCCCTCGGGGGTAT 240
Db 787 GCTTCAACGGGGGCGCAGCAACCACTACTTTGGCTACAGCACCCCTCGGGGGTAT 846
Qy 241 TTTGATTTCAACAGATTCCACTGCCACTTTTCAACAGTACTGGCAGGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCCACTGCCACTTTTCAACAGTACTGGCAGGACTCATCAAC 906
Qy 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCAAGTCAAG 360
Db 907 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCAAGTCAAG 966
Qy 361 GAGGTCACGACGAATGATGGGTCCAAACCACTCGCTTAATAACCTTTACCAGCACGGTTCAA 420
Db 967 GAGGTCACGACGAATGATGGGTCCAAACCACTCGCTTAATAACCTTTACCAGCACGGTTCAA 1026
Qy 421 GTCTTCTCGGACTCGGATACAGATTTCGTCTCGTCTCGGCTCTGGCGACACAGGGCTGC 480
Db 1027 GTCTTCTCGGACTCGGATACAGATTTCGTCTCGTCTCGGCTCTGGCGACACAGGGCTGC 1086
Qy 481 CTCCCTCGGTTCCCGGCGAGCGTGTTCATGTTCCGCAATACGGCTACTCGACGCTCAAC 540
Db 1087 CTCCCTCGGTTCCCGGCGAGCGTGTTCATGTTCCGCAATACGGCTACTCGACGCTCAAC 1146
Qy 541 AATGGCAGCCAGCGGTGGAGCGTTTCATCTTTTACTGCTGGAATATTTCCCTTTCTCAG 600
Db 1147 AATGGCAGCCAGCGGTGGAGCGTTTCATCTTTTACTGCTGGAATATTTCCCTTTCTCAG 1206

QY 601 ATGCTGAGAAAGCGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCTTTCCAC 660
DB 1207 ATGCTGAGAAAGCGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCTTTCCAC 1266
QY 661 AGCAGTACGCGCAGACGAGCCTGGACCGGCTGATGATCTCTCATGACCAATAC 720
DB 1267 AGCAGTACGCGCAGACGAGCCTGGACCGGCTGATGATCTCTCATGACCAATAC 1326
QY 721 CTGTATTACTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGACTTGGCTG 780
DB 1327 CTGTATTACCGNACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGACTTGGCTG 1386
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAATCTGGCTTACCTGACCC 840
DB 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAATCTGGCTTACCTGACCC 1446
QY 841 TGTATTGCGCAGCAGCGGCTTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 900
DB 1447 TGTATTGCGCAGCAGCGGCTTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 1506
QY 901 TGGACTGGTGTCTCAAAATATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960
DB 1507 TGGACTGGTGTCTCAAAATATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 1566
QY 961 GCTATGGCCTCACACAAAGACGACGAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1020
DB 1567 GCTATGGCCTCACACAAAGACGAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1626
QY 1021 TTTGGAAGAGAGCGCGGAGCTTCAACACTGCAATTTGGAACAATGATCATGATACAGAC 1080
DB 1627 TTTGGAAGAGAGCGCGGAGCTTCAACACTGCAATTTGGAACAATGATCATGATACAGAC 1686
QY 1081 GAAGAGGAATTAAGCCACTTAACCTGTGGCCACCGAAGATTTCGGACCGTGGCAGTC 1140
DB 1687 GAAGAGGAATTAAGCCACTTAACCTGTGGCCACCGAAGATTTCGGACCGTGGCAGTC 1746
QY 1141 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGCTATGCGAGCATTA 1200
DB 1747 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGCTATGCGGAGCATTA 1806
QY 1201 CTGGCATGGTGTGGCAAGATAGAGACGTGTACTGAGGGTCCCATTTTGGGCCAAAATT 1260
DB 1807 CTGGCATGGTGTGGCAAGATAGAGACGTGTACTGAGGGTCCCATTTTGGGCCAAAATT 1866
QY 1261 CCTCACAGATGGACACTTTACCGCTCTCTTATGCGCGCTTTGGACCTCAAGAAC 1320
DB 1867 CCTCACAGATGGACACTTTACCGCTCTCTTATGCGCGCTTTGGACCTCAAGAAC 1926
QY 1321 CCGCCTCTCAGATCCTCATCAAAACACACGCTGTCTCTCGAATCCTCCGCGGAGTTT 1380
DB 1927 CCGCCTCTCAGATCCTCATCAAAACACACGCTGTCTCTCGAATCCTCCGCGGAGTTT 1986
QY 1381 TCAGCTACAAAGTTTGCTTCATTCATCCCAATCTCCACAGGACAAGTGAGTGGAA 1440
DB 1987 TCAGCTACAAAGTTTGCTTCATTCATTCATCCCAATCTCCACAGGACAAGTGAGTGGAA 2046
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAGTGCAGTACACA 1500
DB 2047 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAGTGCAGTACACA 2106
QY 1501 TCCAAATTATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACAAATGACTTTTACT 1560
DB 2107 TCCAAATTATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACAAATGACTTTTACT 2166
QY 1561 GAGCCTCGCCCCATTTGGACCCGTTTACCTTTACCCGTCCTCTGTA 1604
DB 2167 GAGCCTCGCCCCATTTGGACCCGTTTACCTTTACCCGTCCTCTGTA 2210

RESULT 14
ADZ26930
ID
ADZ26930 standard; DNA; 2214 BP.
XX

AC ADZ26930;
XX 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus DNA SEQ ID NO 80.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
XX WO2005033321-A2.
XX
XX 14-APR-2005.
XX
XX 30-SEP-2004; 2004WO-US028817.
XX
XX 30-SEP-2003; 2003US-0508226P.
XX
XX 29-APR-2004; 2004US-0566546P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX
XX Claim 50; SEQ ID NO 80; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.
XX
XX
SQ Sequence 2214 BP; 572 A; 658 C; 543 G; 441 T; 0 U; 0 Other;
Query Match 99.4%; Score 1594.4; DB 14; Length 2214;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1598; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATGGCTTTCAGGGCGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGAGTGGGT 60
DB 610 ATGGCTGCAGGGCGTGGCGCTCCAAATGGCAGACAATAACGAAGCGCCGACGAGTGGGT 669
QY 61 AATGCCTCAGAAATTTGGCAATTCGGAATTCACATGCTGGCGCAGACAGTATCACCACC 120
DB 670 AATGCCTCAGAAATTTGGCAATTCGGAATTCACATGCTGGCGCAGACAGTATCACCACC 729
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACAATTAACACCTCTACAAGCAAAATCTCCAGT 180
DB 730 AGCACCCGACCTGGGCGCTTGGCCACCTACAATTAACACCTCTACAAGCAAAATCTCCAGT 789
QY 181 GCTTCAACCGGGGCCAGCAACCAACCACTACTCTCGGCTACAGCAACCCCTCGGGGTAT 240
DB 790 GCTTCAACCGGGGCCAGCAACCAACCACTACTCTCGGCTACAGCAACCCCTCGGGGTAT 849
QY 241 TTTGATTTCAACAGATTCCACTGCCACTTTTCAACCGTACTGGCAGGACTCATCAAC 300
DB 850 TTTGATTTCAACAGATTCCACTGCCACTTTTCAACCGTACTGGCAGGACTCATCAAC 909

QY 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAGTCAAG 360
Db 910 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAGTCAAG 969
QY 361 GAGGTTCACAGCAAAATGATGGCGTCACAAACATCGCTTAATAACCTTACACGACCGGTTCAA 420
Db 970 GAGGTTCACAGCAAAATGATGGCGTCACAAACATCGCTTAATAACCTTACACGACCGGTTCAA 1029
QY 421 GTCTTCTCGGACTCGGAGTACAGACTTCCGTAAGTTCCTCGGTCTGCGGACCAAGGGCTGC 480
Db 1030 GTCTTCTCGGACTCGGAGTACAGACTTCCGTAAGTTCCTCGGTCTGCGGACCAAGGGCTGC 1089
QY 481 CTCCTCTCGGTTCGGCGGAGCGGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Db 1090 CTCCTCTCGGTTCGGCGGAGCGGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 1149
QY 541 AATGGCAGCAAGCCGTGGGAGGTTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 600
Db 1150 AATGGCAGCAAGCCGTGGGAGGTTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 1209
QY 601 ATGCTGAGAACGGGCAACATTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 660
Db 1210 ATGCTGAGAACGGGCAACATTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 1269
QY 661 AGCAGCTACGCGACAGCGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 1270 AGCAGCTACGCGACAGCGAGCCTGGACCGGCTGATGAATCCTCTCATCGTCCCAATAC 1329
QY 721 CTGTATTACTGTAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAGGACTTGTCTG 780
Db 1330 CTGTATTACTGTAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAGGACTTGTCTG 1389
QY 781 TTATAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACATGGCTACTCGACCC 840
Db 1390 TTATAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACATGGCTACTCGACCC 1449
QY 841 TGTATTCCGACAGCGCGCTTTCTAAACAAAACAGACAGCAACAAACAGCAATTTTACC 900
Db 1450 TGTATTCCGACAGCGCGCTTTCTAAACAAAACAGACAGCAACAAACAGCAATTTTACC 1509
QY 901 TGGACTGGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAATCAACCCCTGGCACT 960
Db 1510 TGGACTGGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAATCAACCCCTGGCACT 1569
QY 961 GCTATGGCTTCACACAAAGACGACGAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1020
Db 1570 GCTATGGCTTCACACAAAGACGACGAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1629
QY 1021 TTTGGAAAAGAGCGCGGAGCTTCAAACTGCAATGGCAATGGACATGTCATGATTCAGAC 1080
Db 1630 TTTGGAAAAGAGCGCGGAGCTTCAAACTGCAATGGCAATGGCAATGTCATGATTCAGAC 1689
QY 1081 GAAGAGAAATTAAGCCACTAACCTCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTCT 1140
Db 1690 GAAGAGAAATTAAGCCACTAACCTCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTCT 1749
QY 1141 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGTCATGTCATGGAGCATTA 1200
Db 1750 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGTCATGTCATGGAGCATTA 1809
QY 1201 CTTGGCATGGTGTGGCAAGTAGAGCGTGTACTTCGCGGTCCTCATTTGGGCCAAAATT 1260
Db 1810 CTTGGCATGGTGTGGCAAGTAGAGCGTGTACTTCGCGGTCCTCATTTGGGCCAAAATT 1869
QY 1261 CCTCACACAGATGACACTTTTACCCGCTCTCTCTTATGGCGGCTTTGAGCTCAAGAAC 1320
Db 1870 CCTCACACAGATGACACTTTTACCCGCTCTCTCTTATGGCGGCTTTGAGCTCAAGAAC 1929
QY 1321 CCGCCTCTCTCAGATCTCTCATCAAAAACAGCGCTGTTCTCTCGAATCTCTCGGCGGAGTTT 1380
Db 1930 CCGCCTCTCTCAGATCTCTCATCAAAAACAGCGCTGTTCTCTCGAATCTCTCGGCGGAGTTT 1989

QY 1381 TCAGCTACAAAGTTTGTCTTCAATCACCCTTCAATCACCCTTCAATCACCCTTCAATCACCCTTCAAT 1440
Db 1990 TCAGCTACAAAGTTTGTCTTCAATCACCCTTCAATCACCCTTCAATCACCCTTCAATCACCCTTCAAT 2049
QY 1441 ATTGAATGGGAGCTGCAGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA 1500
Db 2050 ATTGAATGGGAGCTGCAGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA 2109
QY 1501 TCCAAATTATGCAAAATCTGCCAAGTTCATTTTACTGTGGACAAACAATGGACTTTTACT 1560
Db 2110 TCCAAATTATGCAAAATCTGCCAAGTTCATTTTACTGTGGACAAACAATGGACTTTTACT 2169
QY 1561 GAGCCTCGCCCTTGGCAGCCGTTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 1604
Db 2170 GAGCCTCGCCCTTGGCAGCCGTTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 2213
RESULT 15
ADZ26928
ID ADZ26928 standard; DNA; 2211 BP.
XX
AC ADZ26928;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus DNA SEQ ID NO 78.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antitumor; dermatological; ds.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004WO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
PS WPI; 2005-285437/29.
XX
PT New adeno-associated virus (AAV) clade comprising at least three AAV members, useful for preventing and/or treating arthritis, multiple sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial infection and cancer.
XX
PS Claim 19; SEQ ID NO 78; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising at least three AAV members, where each member of the AAV clade is phylogenetically related as determined using a Neighbor-Joining heuristic by a bootstrap value of at least 75 % per 1000 isolates and a Poisson correction distance measurement of no more than 0.05. The methods and/or treatment of rheumatoid arthritis, multiple sclerosis, sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's disease, hemophilia, HIV, bacterial infection, cancer and ulcerative colitis. The present sequence represents an adeno-associated virus DNA.
SQ Sequence 2211 BP; 575 A; 654 C; 541 G; 441 T; 0 U; 0 Other;
Query Match 99.3%; Score 1592.6; DB 14; Length 2211;
Best Local Similarity 99.6%; Pred. No. 0;

Matches 1597; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
Qy	1	ATGGCTTCAGGCGGTGGCGCACCAATATGGCAGACAATTAACGAAGGCGCGCAGCGATGGGT	60
Db	607	ATGGCTTCAGGCGGTGGCGCACCAATATGGCAGACAATTAACGAAGGCGCGCAGCGATGGGT	666
Qy	61	AATGCTCAGGAAATTTGGCAATTTGCATTTCCATGCTGGGCGCAGAGTCAATCAACCACC	120
Db	667	AATGCTCAGGAAATTTGGCAATTTGCATTTCCATGCTGGGCGCAGAGTCAATCAACCACC	726
Qy	121	AGCACCCGACCTGGGCGCTTGGCCACCTACATTAACCACTCTACAGCAAAATCTCCAGT	180
Db	727	AGCACCCGACCTGGGCGCTTGGCCACCTACATTAACCACTCTACAGCAAAATCTCCAGT	786
Qy	181	GCTTCAACGGGGGGCCAGCAACACACCACTACTTTCGGCTACAGACCCCTTGGGGGTAT	240
Db	787	ACTTCAACGGGGGGCCAGCAACACCACTACTTTCGGCTACAGACCCCTTGGGGGTAT	846
Qy	241	TTTGATTTCAACAGATTTCCATGCTGCCATTTTCAACAAGTGAATGGCAGCGACTCATCAAC	300
Db	847	TTTGATTTCAACAGATTTCCATGCTGCCATTTTCAACAAGTGAATGGCAGCGACTCATCAAC	906
Qy	301	AACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAG	360
Db	907	AACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCGAG	966
Qy	361	GAGGTCACGACCAATGATGGCGTCAACAACATCGCTAATAACCTTACAGCAGCGTTCAA	420
Db	967	GAGGTCACGACCAATGATGGCGTCAACAACATCGCTAATAACCTTACAGCAGCGTTCAA	1026
Qy	421	GTCTTCTCGGACTCGGAGTACAGCTTCCGTCTCGGTCTGCGCACCAAGGCTGC	480
Db	1027	GTCTTCTCGGACTCGGAGTACAGCTTCCGTCTCGGTCTGCGCACCAAGGCTGC	1086
Qy	481	CTCCCTTCGTTCCGGCGGACGTTTCATGATTCGCAATACGCTACCTGACGCTCAAC	540
Db	1087	CTCCCTTCGTTCCGGCGGACGTTTCATGATTCGCAATACGCTACCTGACGCTCAAC	1146
Qy	541	AATGCGAGCAAGCGTGGGAGCTTCATCTTTACTGCTGGAATATTTTCCCTTCTCAG	600
Db	1147	AATGCGAGCAAGCGTGGGAGCTTCATCTTTACTGCTGGAATATTTTCCCTTCTCAG	1206
Qy	601	ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTTCCAC	660
Db	1207	ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTTCCAC	1266
Qy	661	AGCAGCTACGGGACAGCCGAGCGCTGGACCGGTGATGAATCTCTCATCGACCAATAC	720
Db	1267	AGCAGCTACGGGACAGCCGAGCGCTGGACCGGTGATGAATCTCTCATCGACCAATAC	1326
Qy	721	CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGACTTTGCTG	780
Db	1327	CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGACTTTGCTG	1386
Qy	781	TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACCTGGACCC	840
Db	1387	TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACCTGGACCC	1446
Qy	841	TGTTATCGGAGAGCGCGTTTCTAAACAAACAAAGACAAACAAACAGCAATTTTACC	900
Db	1447	TGTTATCGGAGAGCGCGTTTCTAAACAAACAAAGACAAACAAACAGCAATTTTACC	1506
Qy	901	TGGACTGGTCTTCAAAATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT	960
Db	1507	TGGACTGGTCTTCAAAATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT	1566
Qy	961	GCTATGGCTCACACAAAGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT	1020
Db	1567	GCTATGGCTCACACAAAGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT	1626
Qy	1021	TTTGGAAAGAGAGCGCCGAGCTTCAAAACATCTGCAATGGACAAATGTCATGATACAGAC	1080
Db	1627	TTTGGAAAGAGAGCGCCGAGCTTCAAGACATGTCATTTGGACAAATGTCATGATACAGAC	1686

Search completed: November 28, 2005, 19:55:47
Job time : 754.542 secs

Qy	1081	GAAGAGGAAATTAAGCCACCTAACCCCTGTGGCCACCAGAAAGATTTTGGGACCGTGGCAGTTC	1140
Db	1687	GAAGAGGAAATTAAGCCACCTAACCCCTGTGGCCACCAGAAAGATTTTGGGACCGTGGCAGTTC	1746
Qy	1141	AATTTCCAGAGCAGCAGCAGACCTTCGGACCGGAGATGTGCATGCTATGGGAGCATTTA	1200
Db	1747	AATTTCCAGAGCAGCAGCAGACCTTCGGACCGGAGATGTGCATGCTATGGGAGCATTTA	1806
Qy	1201	CCTGCAATGGTGTGGCAAGATAGAGACGTGTACCTTGCAGGGTCCCATTTTGGGCCAAAATT	1260
Db	1807	CCTGCAATGGTGTGGCAAGATAGAGACGTGTACCTTGCAGGGTCCCATTTTGGGCCAAAATT	1866
Qy	1261	CCTCACACAGATGGACACTTTTCAACCCGTCTCTCTTATGGGGCGCTTTGGACTCAAGAAC	1320
Db	1867	CCTCACACAGATGGACACTTTTCAACCCGTCTCTCTTATGGGGCGCTTTGGACTCAAGAAC	1926
Qy	1321	CGGCTCTCTCAGATCCTCATCAAAACACGCTGTTCCTGCGAATCCTCCGGCGGAGTTT	1380
Db	1927	CGGCTCTCTCAGATCCTCATCAAAACACGCTGTTCCTGCGAATCCTCCGGCGGAGTTT	1986
Qy	1381	TCAGCTACAAAGTTTGTCTTCAATCACCCTCAATCTCCACAGGACAAAGTGTGGAA	1440
Db	1987	TCAGCTACAAAGTTTGTCTTCAATCACCCTCAATCTCCACAGGACAAAGTGTGGAA	2046
Qy	1441	ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA	1500
Db	2047	ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA	2106
Qy	1501	TCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGCAACAATGCGACTTTATCT	1560
Db	2107	TCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGCAACAATGCGACTTTATCT	2166
Qy	1561	GAGCCTCGCCCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA	1604
Db	2167	GAGCCTCGCCCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA	2210

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 5861.51 Seconds
(without alignments)
15555.198 Million cell updates/sec

Title: US-10-696-900-1_COPY_2829_4432
Perfect score: 1604
Sequence: 1 atggcttcaggcggtggcgc.....tacctaccgctccctgtga 1604

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :
1: gb_ba :
2: gb_in :
3: gb_env :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pr :
9: gb_ro :
10: gb_sts :
11: gb_sy :
12: gb_un :
13: gb_vi :
14: gb_hq :
15: gb_pl :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1604	100.0	1605	6	BD242773	Adeno-ass
2	1604	100.0	1605	6	AR562505	Sequence
3	1604	100.0	1800	6	BD242772	Adeno-ass
4	1604	100.0	1800	6	AR562504	Sequence
5	1604	100.0	2211	6	BD242771	Adeno-ass
6	1604	100.0	2211	6	CS073491	Sequence
7	1604	100.0	2211	6	CS073614	Sequence
8	1604	100.0	2211	6	AR562503	Sequence
9	1604	100.0	4718	6	BD242766	Adeno-ass
10	1604	100.0	4718	6	AR562498	Sequence
11	1604	100.0	4718	6	AX753251	Sequence
12	1604	100.0	4718	13	AF063497	Adeno-ass
13	1604	100.0	7447	6	AR527492	Sequence
14	1600.8	99.8	2211	6	CS073493	Sequence
15	1600.8	99.8	2211	13	AY530607	Adeno-ass
16	1594.4	99.4	2214	6	CS073492	Sequence
17	1594.4	99.4	2214	13	AY530606	Adeno-ass
18	1592.8	99.3	2211	6	CS073490	Sequence

19	1592.8	99.3	2211	13	AY530611	Adeno-ass
20	1591.2	99.2	2211	6	CS073494	Sequence
21	1591.2	99.2	2211	13	AY530609	Adeno-ass
22	1512.8	94.3	2211	6	CS073615	Sequence
23	1512.8	94.3	4683	6	BD242775	Adeno-ass
24	1512.8	94.3	4683	6	AR562507	Sequence
25	1512.8	94.3	4683	13	AF028704	Adeno-ass
26	1104.6	68.9	2214	6	CS073452	Sequence
27	1104.6	68.9	2214	13	AY530567	Adeno-ass
28	1084	67.6	2214	6	CS073453	Sequence
29	1079	67.3	2214	6	CS073448	Sequence
30	1077.4	67.2	2214	6	CS073447	Sequence
31	1075.8	67.1	2217	6	CS073454	Sequence
32	1074.2	67.0	2214	6	CS073444	Sequence
33	1074.2	67.0	2214	6	CS073446	Sequence
34	1074.2	67.0	2214	13	AY530561	Adeno-ass
35	1072.6	66.9	2214	6	CS073451	Sequence
36	1071.6	66.8	2190	13	AY243000	Non-human
37	1071.6	66.8	3094	6	AX753267	Sequence
38	1071	66.8	2214	6	CS073445	Sequence
39	1071	66.8	2214	6	CS073449	Sequence
40	1071	66.8	2214	13	AY530568	Adeno-ass
41	1071	66.8	2214	13	AY530573	Adeno-ass
42	1070	66.7	2190	13	AY242998	Non-human
43	1070	66.7	2190	13	AY242999	Non-human
44	1070	66.7	2214	6	CS073625	Sequence
45	1070	66.7	3095	6	AX753268	Sequence

ALIGNMENTS

RESULT 1
BD242773 1605 bp DNA linear PAT 17-JUL-2003
LOCUS Adeno-associated virus serum type 1 nucleic acid sequence, vector
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector
ACCESSION BD242773.1 GI:33052543
VERSION BD242773.1
KEYWORDS JP 2002529098-A/8.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1605)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector
JOURNAL and host cell containing the same
Patent: JP 2002529098-A 8 10-SEP-2002.
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT OS AAV-1
PN JP 2002529098-A/8
PD 10-SEP-2002
PF 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES M WILSON, WEIDONG XIAO
PC C12N15/09, A61K31/711, A61P43/00, A61P43/00, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02//A61K35/76, C12N15/00, C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH Key Location/Qualifiers
FT CDS (1)..(1602).

FEATURES
source 1..1605
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 1604; DB 6; Length 1605;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 AGCACCCGCACTCGGCGCTTGCCACCTCAATAAACCACTCTACAAAGCAAAATCTCCAGT 180
Qy 181 GCTTCAACGGGGCCAGCAACACACCACTACTCTCGGCTCAGACACCCCTCGGGGTAT 240
Db 181 GCTTCAACGGGGCCAGCAACACCACTACTCTCGGCTCAGACACCCCTCGGGGTAT 240
Qy 241 TTGTATTTCAACAGATTCACCTGCCACTTTTCAACACGTGACTGGCAGCGACTCATCAAC 300
Db 241 TTGTATTTCAACAGATTCACCTGCCACTTTTCAACACGTGACTGGCAGCGACTCATCAAC 300
Qy 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360
Db 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360
Qy 361 GAGGTCAACGAAATGATGGCGTCAACACCATCGCTAATAACCTTACACGACCGGTTCAA 420
Db 361 GAGGTCAACGAAATGATGGCGTCAACACCATCGCTAATAACCTTACACGACCGGTTCAA 420
Qy 421 GTCTTCTCGGACTCGGAGTACCAAGTTCGGTACGTCTCGGCTCTGGCAACGAGGCTGC 480
Db 421 GTCTTCTCGGACTCGGAGTACCAAGTTCGGTACGTCTCGGCTCTGGCAACGAGGCTGC 480
Qy 481 CTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Db 481 CTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Qy 541 AATGGCAGCAACGCGTGGAGCTTCATCTTTTACTGTGCTGGAATATTTCCCTTCTCAG 600
Db 541 AATGGCAGCAACGCGTGGAGCTTCATCTTTTACTGTGCTGGAATATTTCCCTTCTCAG 600
Qy 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660
Db 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660
Qy 661 AGCAGTACGCGCACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 661 AGCAGTACGCGCACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Qy 721 CTGTATTACTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAGGACTTGTCTG 780
Db 721 CTGTATTACTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAGGACTTGTCTG 780
Qy 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAACTGGCTTACCTGACCC 840
Db 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAACTGGCTTACCTGACCC 840
Qy 841 TGTATTGCGGAGCAGCGCGTCTTCAAAACAAACAGACAAACACAGCAATTTTACC 900
Db 841 TGTATTGCGGAGCAGCGCGTCTTCAAAACAAACAGACAAACACAGCAATTTTACC 900
Qy 901 TGGACTGGTGTCTCAAAATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960
Db 901 TGGACTGGTGTCTCAAAATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960
Qy 961 GCTATGGCTCTCACAAAGACGACGAGCAAGTCTTTTCCCATGAGCGGTGTCAATGAT 1020
Db 961 GCTATGGCTCTCACAAAGACGACGAGCAAGTCTTTTCCCATGAGCGGTGTCAATGAT 1020
Qy 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGACAAATGTCAATGATACAGAC 1080
Db 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGACAAATGTCAATGATACAGAC 1080
Qy 1081 GAAGAGGAATTAAGCCACTTAACCTGTGGCCACCGAAGATTTGGACCGTGGCAGTC 1140
Db 1081 GAAGAGGAATTAAGCCACTTAACCTGTGGCCACCGAAGATTTGGACCGTGGCAGTC 1140
Qy 1141 AATTTCCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1200
Db 1141 AATTTCCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1200
Qy 1201 CCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGTGAGGGTCCCATTTGGGCCAAAT 1260
Db 1201 CCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGTGAGGGTCCCATTTGGGCCAAAT 1260

Qy 1261 CCTCACAGATGGACACTTTTACCCGCTCTCTCTTATATGGGCGGCTTTGGACTCAAGAAC 1320
Db 1261 CCTCACAGATGGACACTTTTACCCGCTCTCTCTTATATGGGCGGCTTTGGACTCAAGAAC 1320
Qy 1321 CGGCTCTCTCAGATCCTCATCAAAACACGCGTGTCTTCCGGAATCTTCCGCGGAGTTT 1380
Db 1321 CGGCTCTCTCAGATCCTCATCAAAACACGCGTGTCTTCCGGAATCTTCCGCGGAGTTT 1380
Qy 1381 TCAGCTACAAAGCTTTGCTTTCATTCATCACCCTCAATACTCCACAGGACAAGTGTGGAA 1440
Db 1381 TCAGCTACAAAGCTTTGCTTTCATTCATCACCCTCAATACTCCACAGGACAAGTGTGGAA 1440
Qy 1441 ATTGAATGGAGCTCGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 1441 ATTGAATGGAGCTCGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Qy 1501 TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACATGACTTTTATCT 1560
Db 1501 TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACATGACTTTTATCT 1560
Qy 1561 GAGCCTCGCCCAATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA 1604
Db 1561 GAGCCTCGCCCAATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA 1604

RESULT 3
BD242772 1800 bp DNA linear PAT 17-JUL-2003
LOCUS Adeno-associated virus serum type 1 nucleic acid sequence, vector
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector
ACCESSION BD242772
VERSION BD242772.1 GI:33052542
KEYWORDS JP 2002529098-A/7.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1800)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same
JOURNAL Patent: JP 2002529098-A 7 10-SEP-2002;
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT OS AAV-1
PN JP 2002529098-A/7
PD 10-SEP-2002
PF 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES W WILSON,WEIDONG XIAO
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host

CC cell containing the same
FH Key Location/Qualifiers
FT CDS (1)..(1797).
FEATURES
source 1..1800
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGAGTGGGT 60
Db 196 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGAGTGGGT 255
Qy 61 AATGCTTCAGGAATTTGGCATTTCCACATGGCTGGGCGACAGAGTCATCACCACC 120

```
|||||
256 AATGCTCAGGAAATTGGCAATTGGCATTCACATGGCTGGCGACAGAGTCTATCCACC 315
|||||
121 AGCACCGGACCTGGGCGCTTGGCCACCTTACAAATAACACCTCTCAAGCAAAATCTCCAGT 180
|||||
316 AGCACCGGACCTGGGCGCTTGGCCACCTTACAAATAACACCTCTCAAGCAAAATCTCCAGT 375
|||||
181 GCTTCAAGGGGGCGGACGACCAACACCACTACTTTCGGCTCAGACACCCCTGGGGGTAT 240
|||||
376 GCTTCAAGGGGGCGGACGACCAACCACTACTTTCGGCTCAGACACCCCTGGGGGTAT 435
|||||
241 TTTGATTTCAACAGATTTCCACTGCGCACTTTTACCACGCTGACTGGCAGGACTCATCAAC 300
|||||
436 TTTGATTTCAACAGATTTCCACTGCGCACTTTTACCACGCTGACTGGCAGGACTCATCAAC 495
|||||
301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAAACTCAAGTCAAG 360
|||||
496 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAAACTCAAGTCAAG 555
|||||
361 GAGGTACAGCAAGATGATGGCGTCAACACCATCGCTTAATACTTACCAGACGGTTCAA 420
|||||
556 GAGGTACAGCAAGATGATGGCGTCAACACCATCGCTTAATACTTACCAGACGGTTCAA 615
|||||
421 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTCTCGCAGCACGAGGCTGC 480
|||||
616 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTCTCGCAGCACGAGGCTGC 675
|||||
481 CTCCCTCCGTTCCGGCGGACGTGTTTATGATTCGCAATAACGCTACTGACGCTCAAC 540
|||||
676 CTCCCTCCGTTCCGGCGGACGTGTTTATGATTCGCAATAACGCTACTGACGCTCAAC 735
|||||
541 AATGCGACCAAGCGTGGGAGTTCATCTTTTACTGCGCTGGAATATTTCCCTTCTCAG 600
|||||
736 AATGCGACCAAGCGTGGGAGTTCATCTTTTACTGCGCTGGAATATTTCCCTTCTCAG 795
|||||
601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC 660
|||||
796 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC 855
|||||
661 AGCAGTACGCGACAGCAGAGCTTGACCGGCTGATGAATCTCTCATACGACCAATAC 720
|||||
856 AGCAGTACGCGACAGCAGAGCTTGACCGGCTGATGAATCTCTCATACGACCAATAC 915
|||||
721 CTGTATTACCTGAACAGACTCAAAATCAGTCGGGAAGTGCCTTCAAAACAGGACTTGCTG 780
|||||
916 CTGTATTACCTGAACAGACTCAAAATCAGTCGGGAAGTGCCTTCAAAACAGGACTTGCTG 975
|||||
781 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACTCGACCC 840
|||||
976 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACTCGACCC 1035
|||||
841 TGTATTCCGAGAGCGCGCTTTTAAAAAAGGACAAAGGACAAAGGACAAAGGACAAAGGAC 900
|||||
1036 TGTATTCCGAGAGCGCGCTTTTAAAAAAGGACAAAGGACAAAGGACAAAGGACAAAGGAC 1095
|||||
901 TGGACTGGTGGTCTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960
|||||
1096 TGGACTGGTGGTCTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 1155
|||||
961 GCTATGGCTCTCACAAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTCATGATT 1020
|||||
1156 GCTATGGCTCTCACAAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTCATGATT 1215
|||||
1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGTGCATTTGGACAAATGTATATACAGAC 1080
|||||
1216 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGTGCATTTGGACAAATGTATATACAGAC 1275
|||||
1081 GAGAGGAAATTAAGCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
|||||
1276 GAAGAGGAAATTAAGCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1335
|||||
1141 AATTTCCAGAGCAGCAGACAGACCTTGGCAGCGGAGATGTGCATGCTATGGGAGCATTA 1200
|||||
```

```
Db 1336 AATTTCCAGAGCAGCAGCAGACACCTTGGACCGAGATGTGATCTATGGGAGCATTA 1395
Qy 1201 CCTGGCATGTGTGGCAAGATAGAGACGTGTACTCTGACGGTCCCATTTGGGCGCAAAAT 1260
Db 1396 CCTGGCATGTGTGGCAAGATAGAGACGTGTACTCTGACGGTCCCATTTGGGCGCAAAAT 1455
Qy 1261 CCTCACACAGATGGACACTTTTACCCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1456 CCTCACACAGATGGACACTTTTACCCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1515
Qy 1321 CCGCTCTCAGATCTCTCATCAAAAACAACGCTGTCTCTGCGAATCTCTCCGCGGAGTTT 1380
Db 1516 CCGCTCTCAGATCTCTCATCAAAAACAACGCTGTCTCTGCGAATCTCTCCGCGGAGTTT 1575
Qy 1381 TCAGTCAAAAGTTTGTCTTCAATCATCCCAATACTTCCACAGCAAGTGAAGTGTGGAA 1440
Db 1576 TCAGTCAAAAGTTTGTCTTCAATCATCCCAATACTTCCACAGCAAGTGAAGTGTGGAA 1635
Qy 1441 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 1500
Db 1636 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 1695
Qy 1501 TCCAATTATGCAAAATCTGCCAACGTTTACTTTTACTGTGGACAAATGGACTTTATACT 1560
Db 1696 TCCAATTATGCAAAATCTGCCAACGTTTACTTTTACTGTGGACAAATGGACTTTATACT 1755
Qy 1561 GAGCTCTGCCCATTTGGCACCGCTTACTTACCCTGCCCTGTA 1604
Db 1756 GAGCTCTGCCCATTTGGCACCGCTTACTTACCCTGCCCTGTA 1799

RESULT 4
AR562504
LOCUS AR562504 1800 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 14 from patent US 6759237.
ACCESSION AR562504
VERSION AR562504.1 GI:53976570
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1800)
AUTHORS Wilson, J.M. and Xiao, W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same
JOURNAL Patent: US 6759237-A 14 06-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
Location/Qualifiers
source 1..1800
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCGCGAGTGGGT 60
Db 196 ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCGCGAGTGGGT 255
Qy 61 AATGCTCAGGAATTTGGCATTTGGCATTTGGCATTTGGCGGACAGAGTCTATCCACC 120
Db 256 AATGCTCAGGAATTTGGCATTTGGCATTTGGCATTTGGCGGACAGAGTCTATCCACC 315
Qy 121 AGCACCCGACCTGGGCGCTTGGCCACCTTACAAATAACACCTCTCAAGCAAAATCTCCAGT 180
Db 316 AGCACCCGACCTGGGCGCTTGGCCACCTTACAAATAACACCTCTCAAGCAAAATCTCCAGT 375
Qy 181 GCTTCAAGGGGGCGGACGACCAACCACTACTTTCGGCTCAGACACCCCTGGGGGTAT 240
Db 376 GCTTCAAGGGGGCGGACGACCAACCACTACTTTCGGCTCAGACACCCCTGGGGGTAT 435
```


QY 241 TTTGATTTCAACAGATTCCACTGCGCACTTTTTCACCACTGTTTTCAGCGTGA CTGGCAGGACTCATATCAAC 300
Db 436 TTTGATTTCAACAGATTCCACTGCGCACTTTTTCACCACTGTTTTCAGCGTGA CTGGCAGGACTCATATCAAC 495
QY 301 AACAAATGGGGATTCGGGCCCCAGAGACTCAACTTCAAACTCTTCAACAATCCAAGTCAAG 360
Db 496 AACAAATGGGGATTCGGGCCCCAGAGACTCAACTTCAAACTCTTCAACAATCCAAGTCAAG 555
QY 361 GAGGTCAACGACGAATGATGGCGTCAACAACCATCGCTTAATAA CTTTACAGACACGGTTCAA 420
Db 556 GAGGTCAACGACGAATGATGGCGTCAACAACCATCGCTTAATAA CTTTACAGACACGGTTCAA 615
QY 421 GTCTTCTCGGACTCGGAGTACAGAGTTCCGTAGCTCTCGGTCTGCGCACACAGGCTGC 480
Db 616 GTCTTCTCGGACTCGGAGTACAGAGTTCCGTAGCTCTCGGTCTGCGCACACAGGCTGC 675
QY 481 CTCCTCTCGTTCGGCGGACGTTGTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 540
Db 676 CTCCTCTCGTTCGGCGGACGTTGTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 735
QY 541 AATGGCAGCAAGCCGTGGGAGTTTCATCTTTTACTGCTCGAATATTTCCCTTCTCAG 600
Db 736 AATGGCAGCAAGCCGTGGGAGTTTCATCTTTTACTGCTCGAATATTTCCCTTCTCAG 795
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCTTTCCAC 660
Db 796 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCTTTCCAC 855
QY 661 AGCAGTACCGCACAGCAGCAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 856 AGCAGTACCGCACAGCAGCAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 915
QY 721 CTGTATTACTCGAAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAAAGCACTTGCTG 780
Db 916 CTGTATTACTGACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAAAGCACTTGCTG 975
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGTACTCGGACCC 840
Db 976 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGTACTCGGACCC 1035
QY 841 TGTATTGCGCAGCAGCGCGTTCTTAAACAAACAAACAGACAAACAGCAATTTTACC 900
Db 1036 TGTATTGCGCAGCAGCGCGTTCTTAAACAAACAAACAGACAAACAGCAATTTTACC 1095
QY 901 TGGACTGGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960
Db 1096 TGGACTGGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 1155
QY 961 GCTATGGCCTCACACAAAGACGACGAGCAAGTCTTTTCCATGAGCGGTGTCATGATT 1020
Db 1156 GCTATGGCCTCACACAAAGACGACGAGCAAGTCTTTTCCATGAGCGGTGTCATGATT 1215
QY 1021 TTTGAAAAAGAGCGCGGAGCTTCAACACATGCAATTGGACATGATGATGATTAAGAC 1080
Db 1216 TTTGAAAAAGAGCGCGGAGCTTCAACACATGCAATTGGACATGATGATGATTAAGAC 1275
QY 1081 GAAGAGAAATTAAGCCACTTACCTGTGGCCACCGAAGATTTCGGACCGTGGCAGTC 1140
Db 1276 GAAGAGAAATTAAGCCACTTACCTGTGGCCACCGAAGATTTCGGACCGTGGCAGTC 1335
QY 1141 AATTTTCAGAGCAGCAGCAGACCCCTTGGACCGGAGATGTGTCATGTATGGGAGCATTA 1200
Db 1336 AATTTTCAGAGCAGCAGCAGACCCCTTGGACCGGAGATGTGTCATGTATGGGAGCATTA 1395
QY 1201 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAAT 1260
Db 1396 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAAT 1455
QY 1261 CCTCACACAGATCGACACTTTTCAACCGTCTCTCTTATGGCGGCTTTGAGCTCAAGAAC 1320
Db 1456 CCTCACACAGATCGACACTTTTCAACCGTCTCTCTTATGGCGGCTTTGAGCTCAAGAAC 1515
QY 1321 CCGCCTCTCAGATCCTCATCAAAAAACACGCTGTTCCTTCGCAATCCTCCGGCGGAGTTT 1380

Db 1516 CGGCCTCTCAGATCCTCATCAAAAAACGCGCTGTTCCTGCAATCCTCCGGCGGAGTTT 1575
QY 1381 TCAGCTACAAAAGTTTGTCTTCATTATCAACCAATACTCCACAGGACAAAGTGTGGAA 1440
Db 1576 TCAGCTACAAAAGTTTGTCTTCATTATCAACCAATACTCCACAGGACAAAGTGTGGAA 1635
QY 1441 ATTGAATGGGAGCTGCAGAAAGAAAGACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA 1500
Db 1636 ATTGAATGGGAGCTGCAGAAAGAAAGACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA 1695
QY 1501 TCCAAATATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTTTTACT 1560
Db 1696 TCCAAATATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTTTTACT 1755
QY 1561 GAGCCTCCGCCCAATTGGCACCCGTTTACCTTACCCTGCCCTGTA 1604
Db 1756 GAGCCTCCGCCCAATTGGCACCCGTTTACCTTACCCTGCCCTGTA 1799

RESULT 5
BD242771
LOCUS BD242771 2211 bp DNA linear PAT 17-JUL-2003
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector
ACCESSION BD242771
VERSION BD242771.1 GI:33052541
KEYWORDS JP 2002529098-A/6.
SOURCE unidentifed
ORGANISM unidentifed
REFERENCE 1 (bases 1 to 2211)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector
JOURNAL Patent: JP 2002529098-A 6 10-SEP-2002;
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT OS AAV-1
PN JP 2002529098-A/6
PD 10-SEP-2002
PF 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES M WILSON, WEIDONG XIAO
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH Key Location/Qualifiers
FT CDS (1)..(2208).
FEATURES
source
1..2211
/organism="unidentifed"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGGCCGACGAGTGGGT 60
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGGCCGACGAGTGGGT 666
QY 61 AATGCTCTAGGAAATTTGGCATTGCCATTCCCATGGCTGGGCGACAGAGTCAATCACCACC 120
Db 667 AATGCTCTAGGAAATTTGGCATTGCCATTCCCATGGCTGGGCGACAGAGTCAATCACCACC 726
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACATTAACCACTCTTACAGCAAAATCTCCACT 180
Db 727 AGCACCCGACCTGGGCGCTTGGCCACCTACATTAACCACTCTTACAGCAAAATCTCCACT 786

```
QY 181 GCTTCAACGGGGCCAGCAACACCACTACTCTTGGGCTACAGCACCCCTGGGGTAT 240
Db 787 GCTTCAACGGGGCCAGCAACACCACTACTCTTGGGCTACAGCACCCCTGGGGTAT 846
QY 241 TTTGATTTCAACAGATTCACATGCGCACTTTTCCACACGAGTCTGGCAGGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCACATGCGCACTTTTCCACACGAGTCTGGCAGGACTCATCAAC 906
QY 301 AACAAATTGGGGATTCCGGCCCCAAGAGACTCAACTCTTCAAACTCTTCAACATCCAAAGTCAAG 360
Db 907 AACAAATTGGGGATTCCGGCCCCAAGAGACTCAACTCTTCAAACTCTTCAACATCCAAAGTCAAG 966
QY 361 GAGGTCAACAGCAATGATGGCGTCAACCACTCGCTAATAAATCTTACAGCAGCGTTCAA 420
Db 967 GAGGTCAACAGCAATGATGGCGTCAACCACTCGCTAATAAATCTTACAGCAGCGTTCAA 1026
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTCCTCGGCTCTGCGCACCAAGGCTGC 480
Db 1027 GTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTCCTCGGCTCTGCGCACCAAGGCTGC 1086
QY 481 CTCCCTCGGTTCCGGCGGACGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Db 1087 CTCCCTCGGTTCCGGCGGACGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 1146
QY 541 AATGGAGCAACCGGTGGAGCTTCATCTTTTACTGCTGGATATTTCCCTTCTCAG 600
Db 1147 AATGGAGCAACCGGTGGAGCTTCATCTTTTACTGCTGGATATTTCCCTTCTCAG 1206
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTTCCAC 660
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTTCCAC 1266
QY 661 AGCAGCTACGCGCACAGCAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
Db 1267 AGCAGCTACGCGCACAGCAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 1326
QY 721 CTGTATTACTGACAGAACTCAAAATCAGTCGGGAAGTGCCTTCAACAAAGCACTGCTG 780
Db 1327 CTGTATTACTGACAGAACTCAAAATCAGTCGGGAAGTGCCTTCAACAAAGCACTGCTG 1386
QY 781 TTTAGCGGTGGGTCTCCAGTCGATGCTGTTCAGCCCAAAACCTGGCTACTCGACCC 840
Db 1387 TTTAGCGGTGGGTCTCCAGTCGATGCTGTTCAGCCCAAAACCTGGCTACTCGACCC 1446
QY 841 TGTATTGCGCAGCAGCGCGTTTCTAAAAAAGAGCAAGCAACCAACAGCAATTTTACC 900
Db 1447 TGTATTGCGCAGCAGCGCGTTTCTAAAAAAGAGCAAGCAACCAACAGCAATTTTACC 1506
QY 901 TGGACTGGTCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960
Db 1507 TGGACTGGTCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 1566
QY 961 GCTATGGCTTCACACAAAGACAGCAAGCAAGTCTTTTCCATGAGCGGTGTCATGATT 1020
Db 1567 GCTATGGCTTCACACAAAGACAGCAAGCAAGTCTTTTCCATGAGCGGTGTCATGATT 1626
QY 1021 TTTGGAAAAAGAGAGCGCCGAGCTTCAAAACCTGCAATTTGGACAATGTCATGATTACAGAC 1080
Db 1627 TTTGGAAAAAGAGAGCGCCGAGCTTCAAAACCTGCAATTTGGACAATGTCATGATTACAGAC 1686
QY 1081 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
Db 1687 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1746
QY 1141 AATTTCCAGAGCAGCAGCAGACCTTGGACCGGAGATGTGCTATCGGAGCATTTA 1200
Db 1747 AATTTCCAGAGCAGCAGCAGACCTTGGACCGGAGATGTGCTATCGGAGCATTTA 1806
QY 1201 CCTGGCATGGTGGCAAGATAGAGACGTTGATCTCTGACGGGTCCCAATTTGGGCCAAAAAT 1260
Db 1807 CCTGGCATGGTGGCAAGATAGAGACGTTGATCTCTGACGGGTCCCAATTTGGGCCAAAAAT 1866
```

```
QY 1261 CCTCACAGATGGACACTTTTACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1867 CCTCACAGATGGACACTTTTACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1926
QY 1321 CCGCTCTCTCAGATCTCTCATCAAAAAACACGCTGTTCCTGCGAATCTCTCCGCGGAGTTT 1380
Db 1927 CCGCTCTCTCAGATCTCTCATCAAAAAACACGCTGTTCCTGCGAATCTCTCCGCGGAGTTT 1986
QY 1381 TCAGTCAAAAGTTTGGTTTCATTCATCACCCTTCTTCCAGGACAAAGTGTGTGAA 1440
Db 1987 TCAGTCAAAAGTTTGGTTTCATTCATCACCCTTCTTCCAGGACAAAGTGTGTGAA 2046
QY 1441 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAAGTCAGTACACA 1500
Db 2047 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAAGTCAGTACACA 2106
QY 1501 TCCAATTATGCAAAATCTGCCAAGCTTGATTTTACTGTGGAACAACATGGACTTTTACT 1560
Db 2107 TCCAATTATGCAAAATCTGCCAAGCTTGATTTTACTGTGGAACAACATGGACTTTTACT 2166
QY 1561 GAGCTCTGCCCTATGGCACCCTTACCTTACCGTCCCTGTGA 1604
Db 2167 GAGCTCTGCCCTATGGCACCCTTACCTTACCGTCCCTGTGA 2210

RESULT 6
CS073491          2211 bp      DNA      linear      PAT 05-MAY-2005
LOCUS              Sequence 79 from Patent WO200503321.
DEFINITION
ACCESSION          CS073491
VERSION            CS073491.1  GI:63090470
KEYWORDS
SOURCE              .
ORGANISM            unidentified
                    unclassified
                    unclassified.
REFERENCE
1                   Wilson,J.M., Gao,G., Alvira,M.R. and Vandenbergh,L.H.
AUTHORS             Adeno-associated virus (aav) clades, sequences, vectors containing
TITLE               same, and uses therefor
JOURNAL             Patent: WO 200503321-A 79 14-APR-2005;
                    The Trustees of the University of Pennsylvania (US)
FEATURES
                    Location/Qualifiers
                    1..2211
                    /organism="unidentified"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32644"
                    /note="new AAV serotype, clone rh.71"

ORIGIN
Query Match          100.0%; Score 1604; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGCGACAGCAATAACGAAGCGCCGACGAGTGGGT 60
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGCGACAGCAATAACGAAGCGCCGACGAGTGGGT 666
QY 61 AATGCTTCAGGAAATTTGGCATTTGGCATTTCCATGGCTGGGCGACAGAGTCAATCACCACC 120
Db 667 AATGCTTCAGGAAATTTGGCATTTGGCATTTCCATGGCTGGGCGACAGAGTCAATCACCACC 726
QY 121 AGCACCCGACCTTGGCGCTTGGCCACCTTACATAACCACTTCTACAGCAATCTCCAGT 180
Db 727 AGCACCCGACCTTGGCGCTTGGCCACCTTACATAACCACTTCTACAGCAATCTCCAGT 786
QY 181 GCTTCAACGGGGCCAGCAACACCACTACTCTTGGGCTACAGCACCCCTGGGGTAT 240
Db 787 GCTTCAACGGGGCCAGCAACACCACTACTCTTGGGCTACAGCACCCCTGGGGTAT 846
QY 241 TTTGATTTCAACAGATTCACATGCGCACTTTTCCACACGAGTCTGGCAGGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCACATGCGCACTTTTCCACACGAGTCTGGCAGGACTCATCAAC 906
```

QY 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360
Db 907 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 966
QY 361 GAGGTACAGAGAAATGATGGCGTCACAAACATCGCTTAATAACCTTACCAGCAGCGTTCAA 420
Db 967 GAGGTACAGAGAAATGATGGCGTCACAAACATCGCTTAATAACCTTACCAGCAGCGTTCAA 1026
QY 421 GTCTTCTCGAGCTCGAGATACAGCTTCCGTACGTCCTCGGCTCTCGGACACAGGCGTCG 480
Db 1027 GTCTTCTCGAGCTCGAGATACAGCTTCCGTACGTCCTCGGCTCTCGGACACAGGCGTCG 1086
QY 481 CTCCTCTCGGTCCTCGGCGGACGCTGTTATGATTCGGAATACGCTACCTGACGCTCAAC 540
Db 1087 CTCCTCTCGGTCCTCGGCGGACGCTGTTATGATTCGGAATACGCTACCTGACGCTCAAC 1146
QY 541 AATGGCAGCCAAAGCGTGGGACGTTTCATCTTTTACTGCGTGGAAATATTTCCCTTCTCAG 600
Db 1147 AATGGCAGCCAAAGCGTGGGACGTTTCATCTTTTACTGCGTGGAAATATTTCCCTTCTCAG 1206
QY 601 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTTACACCTTTTGAGGAAGTGCCTTTCCAC 660
Db 1207 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTTACACCTTTTGAGGAAGTGCCTTTCCAC 1266
QY 661 AGCAGCTACGCGCACAGCCAGAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
Db 1267 AGCAGCTACGCGCACAGCCAGAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 1326
QY 721 CTGTATTACTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAAACAAAGGACTTGTCTG 780
Db 1327 CTGTATTACTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAAACAAAGGACTTGTCTG 1386
QY 781 TTTAGCGGTGGTCTCCAGCTGCGATGCTGTGTCAGCCCAAAACTGGCTACCTGGACCC 840
Db 1387 TTTAGCGGTGGTCTCCAGCTGCGATGCTGTGTCAGCCCAAAACTGGCTACCTGGACCC 1446
QY 841 TGTATTCCGCGACGACGCGCTTTCTAAACAAAAACAGACAAACAAACGAAATTTTACC 900
Db 1447 TGTATTCCGCGACGACGCGCTTTCTAAACAAAAACAGACAAACAAACGAAATTTTACC 1506
QY 901 TGGACTGGTCTTCAAAATTAACCTCAATGGGGGTGAATCCATCATCAACCCCTGGCACT 960
Db 1507 TGGACTGGTCTTCAAAATTAACCTCAATGGGGGTGAATCCATCATCAACCCCTGGCACT 1566
QY 961 GCTATGCGCTCACAAAGACGAGAGCAAGTCTTCCCATGAGCGGTGTCAATGAT 1020
Db 1567 GCTATGCGCTCACAAAGACGAGAGCAAGTCTTCCCATGAGCGGTGTCAATGAT 1626
QY 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGSCATTGGACAAATGTCATGATTACAGAC 1080
Db 1627 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGSCATTGGACAAATGTCATGATTACAGAC 1686
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGGAGTC 1140
Db 1687 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGGAGTC 1746
QY 1141 AATTTCCAGAGCAGCAGACGACCTTCCGACCGGAGATGTGCATGCTATGGGAGCATTA 1200
Db 1747 AATTTCCAGAGCAGCAGACGACCTTCCGACCGGAGATGTGCATGCTATGGGAGCATTA 1806
QY 1201 CCTGGCATGGTGGCAAGATAGAGAGTGTACCTGACAGGCTCCCATTTTGGGCCAAAAT 1260
Db 1807 CCTGGCATGGTGGCAAGATAGAGAGTGTACCTGACAGGCTCCCATTTTGGGCCAAAAT 1866
QY 1261 CCTCACAGATGGACACTTTTACCCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1867 CCTCACAGATGGACACTTTTACCCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1926
QY 1321 CCGCTCTCAGATCTCATCAAAAAACAGCTGTTCTCTGGGAATCTCTCGGCGGAGTTT 1380
Db 1927 CCGCTCTCAGATCTCATCAAAAAACAGCTGTTCTCTGGGAATCTCTCGGCGGAGTTT 1986
QY 1381 TCAGCTACAAAGTTTGTCTTATTTCATTCACCCCAATACTCCACAGGACAAAGTGAAGTGCGAA 1440

Db 1987 TCAGCTACAAAGTTTGTCTTATTTCATTCACCCCAATACTCCACAGGACAAGTGAAGTGGAA 2046
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 2047 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 2106
QY 1501 TCCAAATTATGCAAAATCTGCCAAAGTTGATTTTACTGTGGACAAACAAATGGACTTTATCT 1560
Db 2107 TCCAAATTATGCAAAATCTGCCAAAGTTGATTTTACTGTGGACAAACAAATGGACTTTATCT 2166
QY 1561 GAGCCTCGCCCATTTGGCACCCGTTACCTTACCCTGCCCTGTA 1604
Db 2167 GAGCCTCGCCCATTTGGCACCCGTTACCTTACCCTGCCCTGTA 2210

RESULT 7
CS073614
LOCUS CS073614 2211 bp DNA linear PAT 05-MAY-2005
DEFINITION Sequence 202 from Patent WO2005033321.
ACCESSION CS073614
VERSION CS073614.1 GI:63090494
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Wilson,J.M., Gao,G., Alvira,M.R. and Vandenberghe,L.H.
TITLE Adeno-associated virus (aav) clades, sequences, vectors containing
JOURNAL same, and uses therefor
PATENT Patent: WO 2005033321-A 202 14-APR-2005;
The Trustees of the University of Pennsylvania (US)
FEATURES
1. .2211
source /organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32844"
/note="adeno-associated virus, serotype 1"

ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGCGCGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGGAGTGGGT 60
Db 607 ATGGCTTCAGCGCGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGGAGTGGGT 666
QY 61 AATGCCTCAGGAAATTTGGCAATTCGGAATTCACATGGCTGGGCGACAGAGTCAATCACCAC 120
Db 667 AATGCCTCAGGAAATTTGGCAATTCGGAATTCACATGGCTGGGCGACAGAGTCAATCACCAC 726
QY 121 AGCACCGGACCTGGGCGCTTGGCCAGCTACAATAACCACTCTACAGCAAAATCTCCAGT 180
Db 727 AGCACCGGACCTGGGCGCTTGGCCAGCTACAATAACCACTCTACAGCAAAATCTCCAGT 786
QY 181 GCTTCAACGGGGCGCAGCAACGACCACTACTCTCGGCTACAGCAACCCCTCGGGGGTAT 240
Db 787 GCTTCAACGGGGCGCAGCAACGACCACTACTCTCGGCTACAGCAACCCCTCGGGGGTAT 846
QY 241 TTTGATTTCAACAGATTCCTACTGCCACTTTTTCACAGTGCAGTGCAGGACGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCCTACTGCCACTTTTTCACAGTGCAGTGCAGGACGACTCATCAAC 906
QY 301 AACAAATTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCCAAGTCAAG 360
Db 907 AACAAATTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCCAAGTCAAG 966
QY 361 GAGGTACAGAGAAATGATGGCGTCACAAACATCGCTTAATAACCTTACCAGCAGCGTTCAA 420
Db 967 GAGGTACAGAGAAATGATGGCGTCACAAACATCGCTTAATAACCTTACCAGCAGCGTTCAA 1026
QY 421 GTCTTCTCGGACTCGGAGTACCAGCTTCCGTAAGTCTCGGCTCTGCGCACAGGCGTGC 480

Db 1027 GTCTTCGGACTCGAGTACAGCTTCGTAAGTCTCGGCTCTGGCGACACAGGCTGC 1086
Qy 481 CTCCTCCGTTCCCGCGGACGTTGATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Db 1087 CTCCTCCGTTCCCGCGGACGTTGATGATTCGCAATACGGCTACCTGACGCTCAAC 1146
Qy 541 AATGGCAGCAGCCGTGGAGCTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 600
Db 1147 AATGGCAGCAGCCGTGGAGCTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 1206
Qy 601 ATGCTGAGAACGGCAACAACTTTACTCTCAGTACACCTTTTGAGGAAGTGCCTTTCCAC 660
Db 1207 ATGCTGAGAACGGCAACAACTTTACTCTCAGTACACCTTTTGAGGAAGTGCCTTTCCAC 1266
Qy 661 AGCAGTACGCGACACAGCAGGCTTGGACCGGCTGATGAATCTCTCTCATGACCAATAC 720
Db 1267 AGCAGTACGCGACACAGCAGGCTTGGACCGGCTGATGAATCTCTCTCATGACCAATAC 1326
Qy 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAGGAAGTGTG 780
Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAGGAAGTGTG 1386
Qy 781 TTTAGCGGTGCTCCAGCTGGCATGCTGTTTTCAGCCCAAAACTGGCTACTCGACCC 840
Db 1387 TTTAGCGGTGCTCCAGCTGGCATGCTGTTTTCAGCCCAAAACTGGCTACTCGACCC 1446
Qy 841 TGTATTCCGACAGCGCGTTCCTTAAACAAAAACAGACAAACAGCAAGCAATTTTACC 900
Db 1447 TGTATTCCGACAGCGCGTTCCTTAAACAAAAACAGACAAACAGCAAGCAATTTTACC 1506
Qy 901 TGGACTGTGCTTCAAAATATTAACCTCAATGGCGGTGAATTCATCATCAACCTGGCACT 960
Db 1507 TGGACTGTGCTTCAAAATATTAACCTCAATGGCGGTGAATTCATCATCAACCTGGCACT 1566
Qy 961 GCTATGGCTCTCACAAAACAGCAGGAAGCAAGTCTTTCCATAGAGCGGTGTCATGATT 1020
Db 1567 GCTATGGCTCTCACAAAACAGCAGGAAGCAAGTCTTTCCATAGAGCGGTGTCATGATT 1626
Qy 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATTTGGACAAATGTCATGATTAACAGAC 1080
Db 1627 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATTTGGACAAATGTCATGATTAACAGAC 1686
Qy 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 1687 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746
Qy 1141 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGCATCTATGGAGCATTA 1200
Db 1747 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGCATCTATGGAGCATTA 1806
Qy 1201 CTGGCATGTTGGCAAGATAGAGCGTGTACTGCAAGGTCCTCAATTTGGGCCAAAAT 1260
Db 1807 CTGGCATGTTGGCAAGATAGAGCGTGTACTGCAAGGTCCTCAATTTGGGCCAAAAT 1866
Qy 1261 CTTCACACAGATGGAACATTTTCAACCGTCTCTTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1867 CTTCACACAGATGGAACATTTTCAACCGTCTCTTTATGGCGGCTTTGGACTCAAGAAC 1926
Qy 1321 CGCCCTCTCAGATCTCATCAAAACACGCTGTTCTTCGGAATCTCCGGCGGAGTTT 1380
Db 1927 CGCCCTCTCAGATCTCATCAAAACACGCTGTTCTTCGGAATCTCCGGCGGAGTTT 1986
Qy 1381 TCAGCTACAAAGTTTGTCTTCAATTCATCACCATACTCCACAGGACAGTGTGGAA 1440
Db 1987 TCAGCTACAAAGTTTGTCTTCAATTCATCACCATACTCCACAGGACAGTGTGGAA 2046
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA 1500
Db 2047 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA 2106
Qy 1501 TCCAAATTATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACAAATGACTTTTACT 1560

Db 2107 TCCAAATTATGCAAAATCTGCCAAGCTTGATTTTACTGTGGACAAATGAGCTTTTACT 2166
Qy 1561 GAGCTCTGCCCATTTGGCACCCGTTACTTTACCGTCCCTGTGA 1604
Db 2167 GAGCTCTGCCCATTTGGCACCCGTTACTTTACCGTCCCTGTGA 2210
RESULT 8
AR562503 AR562503 2211 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 12 from patent US 6759237.
ACCESSION AR562503
VERSION AR562503.1 GI:53976569
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2211)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same
Patent: US 6759237-A 12 06-JUL-2004;
JOURNAL The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
source 1..2211
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 60
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 666
Qy 61 AATGCTCAGGAATTTGGCATTTGGATTTCCATAGCTGGTGGGCGACAGAGTATCAACACC 120
Db 667 AATGCTCAGGAATTTGGCATTTGGATTTCCATAGCTGGTGGGCGACAGAGTATCAACACC 726
Qy 121 AGCACCCGACCTGGGCTTGGCCACCTACAAATTAACCACTCTCAACGAATAATCTCCAGT 180
Db 727 AGCACCCGACCTGGGCTTGGCCACCTACAAATTAACCACTCTCAACGAATAATCTCCAGT 786
Qy 181 GCTTCAACGGGCGCGACGAACCAACCACTACTTTGGCTACAGCACCCCTGGGGGTAT 240
Db 787 GCTTCAACGGGCGCGACGAACCAACCACTACTTTGGCTACAGCACCCCTGGGGGTAT 846
Qy 241 TTTGATTTCAACAGATTTCCACTGGCACTTTTCCACACGTGACTGGCAGGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTTCCACTGGCACTTTTCCACACGTGACTGGCAGGACTCATCAAC 906
Qy 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCCAAGTCAAG 360
Db 907 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCCAAGTCAAG 966
Qy 361 GAGGTCAAGCAAGATGATGGGTGTCACACCAATCGCTTAATAACCTTTACCAGCAAGTTCAA 420
Db 967 GAGGTCAAGCAAGATGATGGGTGTCACACCAATCGCTTAATAACCTTTACCAGCAAGTTCAA 1026
Qy 421 GTCTTCTCGGACTCGGAGTACAGGTTCCGTAAGTCTCGGCTCTGGCGACACAGGCTGC 480
Db 1027 GTCTTCTCGGACTCGGAGTACAGGTTCCGTAAGTCTCGGCTCTGGCGACACAGGCTGC 1086
Qy 481 CTCCTCTCGGCTCCGGCGGAGCTGTTGATGTTCCGCAATACGGCTACTCGACGCTCAAC 540
Db 1087 CTCCTCTCGGCTCCGGCGGAGCTGTTGATGTTCCGCAATACGGCTACTCGACGCTCAAC 1146
Qy 541 AATGGCAGCAGCCGTGGAGCTTCATCTCTTTTACTGCTGGAATAATTTCCCTTCTCAG 600
Db 1147 AATGGCAGCAGCCGTGGAGCTTCATCTCTTTTACTGCTGGAATAATTTCCCTTCTCAG 1206

QY 601 ATGCTGAGAAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGAAAGTGCCTTTCCAC 660
Db 1207 ATGCTGAGAAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGAAAGTGCCTTTCCAC 1266
QY 661 AGCAGTACCGCAGCAGCAGAGCCTGGACCGCTGATGAACTCTCTCATCGACCAATAC 720
Db 1267 AGCAGTACCGCAGCAGCAGAGCCTGGACCGCTGATGAACTCTCTCATCGACCAATAC 1326
QY 721 CTGTATTACTCTGAACAGAACTCAAAATCACTCCGGAAGTGCCTTGAAGCAAGGACTTGTCTG 780
Db 1327 CTGTATTACTCTGAACAGAACTCAAAATCACTCCGGAAGTGCCTTGAAGCAAGGACTTGTCTG 1386
QY 781 TTTAGCGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACTCGACCC 840
Db 1387 TTTAGCGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACTCGACCC 1446
QY 841 TGTATTCCGAGCAGCGCGTCTTCAAAACAAAACAGACAAACAAACAGCAATTTTACC 900
Db 1447 TGTATTCCGAGCAGCGCGTCTTCAAAACAAAACAGACAAACAAACAGCAATTTTACC 1506
QY 901 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCCATCATCAACCCCTGGCACT 960
Db 1507 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCCATCATCAACCCCTGGCACT 1566
QY 961 GCTATGGCTTCACAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTCATGAT 1020
Db 1567 GCTATGGCTTCACAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTCATGAT 1626
QY 1021 TTTGGAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATTACAGAC 1080
Db 1627 TTTGGAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATTACAGAC 1686
QY 1081 GAAGAGAAATTAAGCCACTTAACCTCTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 1687 GAAGAGAAATTAAGCCACTTAACCTCTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746
QY 1141 AATTTCCAGACGACGACAGACCCCTGGACCGGAGATGTGATGCTATGGGAGCATTA 1200
Db 1747 AATTTCCAGACGACGACAGACCCCTGGACCGGAGATGTGATGCTATGGGAGCATTA 1806
QY 1201 CTTGGCATGTTGGCAAGATAGAGAGCTGTACCTGACGGGTCCCATTTTGGGCCAAATTT 1260
Db 1807 CTTGGCATGTTGGCAAGATAGAGAGCTGTACCTGACGGGTCCCATTTTGGGCCAAATTT 1866
QY 1261 CCTCACACAGATGGACATTTTACCCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1867 CCTCACACAGATGGACATTTTACCCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1926
QY 1321 CCGCCTCCTCAGATCCTCATCAAAACACACGCTGTTCCTGCGAATCCTCCGGCGGAGTTT 1380
Db 1927 CCGCCTCCTCAGATCCTCATCAAAACACACGCTGTTCCTGCGAATCCTCCGGCGGAGTTT 1986
QY 1381 TCAGCTACAAAGTTTGTTCATCAACCCAAATATCTCCAAGCAAGTGAAGTGGAA 1440
Db 1987 TCAGCTACAAAGTTTGTTCATCAACCCAAATATCTCCAAGCAAGTGAAGTGGAA 2046
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 2047 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 2106
QY 1501 TCCAAATTATGCAAAATCTGCAACGTTTGAATTTTACTGTGGACAAACATGGACTTTTACT 1560
Db 2107 TCCAAATTATGCAAAATCTGCAACGTTTGAATTTTACTGTGGACAAACATGGACTTTTACT 2166
QY 1561 GAGCCTCGCCCAATTTGACACCGGTACTTTACCGTCCCTCTGTA 1604
Db 2167 GAGCCTCGCCCAATTTGACACCGGTACTTTACCGTCCCTCTGTA 2210

RESULT 9
BD242766
LOCUS
DEFINITION BD242766 4718 bp DNA linear PAT 17-JUL-2003
Adeno-associated virus serum type 1 nucleic acid sequence, vector

and host cell containing the same.
BD242766
VERSION BD242766.1 GI:33052536
KEYWORDS JP 2002529098-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector
JOURNAL Patent: JP 2002529098-A 1 10-SEP-2002;
COMMENT THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
OS AAV-1
PN JP 2002529098-A/1
PD 10-SEP-2002
PF 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES M WILSON,WEIDONG XIAO
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19,PC
C12N1/21,
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH Key Location/Qualifiers
FT CDS (335)..(2206)
FT CDS (2223)..(4430).
FEATURES
source 1..4718
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGCGGTGGCGCACCACCAATGGCAGACAATAACGAAGCGCGCAGGATGGGT 60
Db 2829 ATGGCTTCAGGCGGTGGCGCACCACCAATGGCAGACAATAACGAAGCGCGCAGGATGGGT 2888
QY 61 AATGCCTCAGGAAATTTGGCATTTCCGATTCACATGGCTGGGCGACAGAGTCATCACACC 120
Db 2889 AATGCCTCAGGAAATTTGGCATTTCCGATTCACATGGCTGGGCGACAGAGTCATCACACC 2948
QY 121 AGACCCGACACTGGGCGCTTGGCCACCTCAATAACCACTCTACAAGCAAAATCTCCAGT 180
Db 2949 AGACCCGACACTGGGCGCTTGGCCACCTCAATAACCACTCTACAAGCAAAATCTCCAGT 3008
QY 181 GCTTCAACGGGGCCAGCAACCACTACTTCGGCTACAGACACCCCTCGGGGTAT 240
Db 3009 GCTTCAACGGGGCCAGCAACCACTACTTCGGCTACAGACACCCCTCGGGGTAT 3068
QY 241 TTTGATTTCAACAGATTCCACTGCCACTTTTACCACGTGACTGGCAGCGACTCATCAAC 300
Db 3069 TTTGATTTCAACAGATTCCACTGCCACTTTTACCACGTGACTGGCAGCGACTCATCAAC 3128
QY 301 AACAAATTGGGGAATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360
Db 3129 AACAAATTGGGGAATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 3188
QY 361 GAGGTCAACGACGAATGATGGCGTCACAACCATCGCTAAATACCTTACCAGCAGCGTTCAA 420
Db 3189 GAGGTCAACGACGAATGATGGCGTCACAACCATCGCTAAATACCTTACCAGCAGCGTTCAA 3248
QY 421 GTCTTCTCGGACTCGGAGTACCAAGCTTCCGTACGTCTCGGCTCTCGGACACCGGCTGC 480
Db 3249 GTCTTCTCGGACTCGGAGTACCAAGCTTCCGTACGTCTCGGCTCTCGGACACCGGCTGC 3308
QY 481 CTCCTCTCGGTTCCCGCGGACAGTGTTCATGATTCGGAATACGGCTACCTGAGCGCTCAAC 540

Db	3309	CTCCCTCCGTTCCCGCGACGCTGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC	3368
Qy	541	AATGCGAGCAAGCGTGGAGCGTTTCATCTCTTTTACTGCTGGAATATTTTCCCTTCTCAG	600
Db	3369	AATGCGAGCAAGCGTGGAGCGTTTCATCTCTTTTACTGCTGGAATATTTTCCCTTCTCAG	3428
Qy	601	ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC	660
Db	3429	ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC	3488
Qy	661	AGCAGCTACGCGACACAGCGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC	720
Db	3489	AGCAGCTACGCGACACAGCGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC	3548
Qy	721	CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAAACGAAGACTTGTG	780
Db	3549	CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAAACGAAGACTTGTG	3608
Qy	781	TTTAGCGGTGGTCTCCAGCTGGCATGTCTGTTTCAGGCCCAAAAACCTGGCTACTCGACCC	840
Db	3609	TTTAGCGGTGGTCTCCAGCTGGCATGTCTGTTTCAGGCCCAAAAACCTGGCTACTCGACCC	3668
Qy	841	TGTTATCGCGACGCGCGTTTCTAAAAAAGAGACAGCAAAACAGCAAACTTTTACC	900
Db	3669	TGTTATCGCGACGCGCGTTTCTAAAAAAGAGACAGCAAAACAGCAAACTTTTACC	3728
Qy	901	TGAGCTGTGCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT	960
Db	3729	TGAGCTGTGCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT	3788
Qy	961	GCTATGGCTCTCACAAAAGACAGCAAGTCTTTTCCATGAGCGGTGTCATGATT	1020
Db	3789	GCTATGGCTCTCACAAAAGACAGCAAGTCTTTTCCATGAGCGGTGTCATGATT	3848
Qy	1021	TTTGGAAAAGAGCGCGGAGCTTCAAAACACTGCAATTGGACAAATGTCATGATTACAGAC	1080
Db	3849	TTTGGAAAAGAGCGCGGAGCTTCAAAACACTGCAATTGGACAAATGTCATGATTACAGAC	3908
Qy	1081	GAAGAGAAATTAAGCCACTAACCCCTGTGGCGACCGAAAGATTTTGGACCGTGGCAGTC	1140
Db	3909	GAAGAGAAATTAAGCCACTAACCCCTGTGGCGACCGAAAGATTTTGGACCGTGGCAGTC	3968
Qy	1141	AATTTCCAGACGACGACAGACCTGCGACCGGAGATGTGATGCTATGGAGCATTA	1200
Db	3969	AATTTCCAGACGACGACAGACCTGCGACCGGAGATGTGATGCTATGGAGCATTA	4028
Qy	1201	CCTGGCATGGTGTGGCAAGATAGACGCTGTACCTGACGGTCCCAATTTGGGCCAAAAT	1260
Db	4029	CCTGGCATGGTGTGGCAAGATAGACGCTGTACCTGACGGTCCCAATTTGGGCCAAAAT	4088
Qy	1261	CCTCACAGATGGACACTTTACCCGCTCTCTTATTTGGGGGCTTTGGACTCAAGAAC	1320
Db	4089	CCTCACAGATGGACACTTTACCCGCTCTCTTATTTGGGGGCTTTGGACTCAAGAAC	4148
Qy	1321	CGCGCTCTCAGATCTCATCAAAAACAGCCTGTTCCTGCGAATCCTCCGGCGAGTTT	1380
Db	4149	CGCGCTCTCAGATCTCATCAAAAACAGCCTGTTCCTGCGAATCCTCCGGCGAGTTT	4208
Qy	1381	TCAGCTACAAAGTTTGCTTTTCATTTCATCACCCCAATCTCCACAGGACAGTGTGGAA	1440
Db	4209	TCAGCTACAAAGTTTGCTTTTCATTTCATCACCCCAATCTCCACAGGACAGTGTGGAA	4268
Qy	1441	ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA	1500
Db	4269	ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA	4328
Qy	1501	TCCAAATTATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGGACTTTTACT	1560
Db	4329	TCCAAATTATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGGACTTTTACT	4388
Qy	1561	GAGCCTCGCCCCATTTGGCACCCGTTTACCTTTACCCGTCCTCTGTA	1604
Db	4389	GAGCCTCGCCCCATTTGGCACCCGTTTACCTTTACCCGTCCTCTGTA	4432
<hr/>			
RESULT 10			
AR562498			
LOCUS			
DEFINITION			
AR562498			
ACCESSION			
AR562498.1			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1..4718			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1604; Conservative			
100.0%; Score 1604; DB 6; Length 4718;			
100.0%; Pred. No. 0;			
0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGCAATAAAGAGGCGCGACGGAGTGGGT	60
Db	2829	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGCAATAAAGAGGCGCGACGGAGTGGGT	2888
Qy	61	AATGCTCAGGAAATTTGGCATTTCCATGCTGGCGCAGAGTCAATCACCACC	120
Db	2889	AATGCTCAGGAAATTTGGCATTTCCATGCTGGCGCAGAGTCAATCACCACC	2948
Qy	121	AGCACCCGACCTGGCGCTTGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT	180
Db	2949	AGCACCCGACCTGGCGCTTGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT	3008
Qy	181	GCTTCAACGGGGCGCGACCAACCACTACTTTCGGCTCAGCACCCCTCTGGGGGTAT	240
Db	3009	GCTTCAACGGGGCGCGACCAACCACTACTTTCGGCTCAGCACCCCTCTGGGGGTAT	3068
Qy	241	TTTGATTTCAACAGATTTCACTGCGACCTTTTCCACAGCTGAGTGGCAGGACTCATCAAC	300
Db	3069	TTTGATTTCAACAGATTTCACTGCGACCTTTTCCACAGCTGAGTGGCAGGACTCATCAAC	3128
Qy	301	AACAATTTGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAATCCCAAGTCAAG	360
Db	3129	AACAATTTGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAATCCCAAGTCAAG	3188
Qy	361	GAGGTCAAGCAAGATGATGGCGTCAACCACTCGCTTAATAAATTTACCAAGCAAGTTCAA	420
Db	3189	GAGGTCAAGCAAGATGATGGCGTCAACCACTCGCTTAATAAATTTACCAAGCAAGTTCAA	3248
Qy	421	GTCTTCTGGACTCGAGTACAGCTTCGGTACGCTCTCGGCTCTGGCGACAGGCTGC	480
Db	3249	GTCTTCTGGACTCGAGTACAGCTTCGGTACGCTCTCGGCTCTGGCGACAGGCTGC	3308
Qy	481	CTCCCTCCGTTCCCGGCGAGCTGTTGATGTTCCGCAATACGGCTACCTGACGCTCAAC	540
Db	3309	CTCCCTCCGTTCCCGGCGAGCTGTTGATGTTCCGCAATACGGCTACCTGACGCTCAAC	3368
Qy	541	AATGCGACCAAGCGGTGGAGCTTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTCAG	600
Db	3369	AATGCGACCAAGCGGTGGAGCTTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTCAG	3428
Qy	601	ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC	660
Db	3429	ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC	3488
Qy	661	AGCAGCTACGCGACACAGCCAGGCGCTGGAGCCGCTGATGAATCCTCTCATCGACCAATAC	720

|||||
3489 AGCAGCTACGCGCACAGCGCCTGGAGCCGGCTGATGAATCCTCTCATCGACCAATAC 3548
721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 780
3549 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 3608
781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGTGAACC 840
3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGTGAACC 3668
841 TGTATTCCGAGCAGCGCGTCTTCAAAACAAAAACAGACAAACAAACAGCAATTTTACC 900
3669 TGTATTCCGAGCAGCGCGTCTTCAAAACAAAAACAGACAAACAAACAGCAATTTTACC 3728
901 TGGACTGGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960
3729 TGGACTGGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 3788
961 GCTATGGCTTCACACAAAGACGACGAAGACAAGTCTTCTTCCCATGAGCGGTGTCTATGATT 1020
3789 GCTATGGCTTCACACAAAGACGACGAAGACAAGTCTTCTTCCCATGAGCGGTGTCTATGATT 3848
1021 TTTGGAAGAGAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAATGTCTATGATTACAGAC 1080
3849 TTTGGAAGAGAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAATGTCTATGATTACAGAC 3908
1081 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTTC 1140
3909 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTTC 3968
1141 AATTTCCAGAGCAGCAGCAGACCTCTGACCGGAGATGTGCTATGGGAGCATTTA 1200
3969 AATTTCCAGAGCAGCAGCAGACCTCTGACCGGAGATGTGCTATGGGAGCATTTA 4028
1201 CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTTGGGCCAAATTT 4088
1261 CCTCACAGATGGACATTTACCCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
4089 CCTCACAGATGGACATTTTACCCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 4148
1321 CGCCCTCTCAGATCTCATCAAAAACAGCCCTGTTCTTCGGAATCCTCCGGCGGAGTTT 1380
4149 CCBCCTCTCAGATCTCTCAAAAACAGCCCTGTTCTTCGGAATCCTCCGGCGGAGTTT 4208
1381 TCAGCTACAAAGTTTGGTTCATTCATCACCCTTCTTCCAGGACCAAGTGTGGAA 1440
4209 TCAGCTACAAAGTTTGGTTCATTCATCACCCTTCTTCCAGGACCAAGTGTGGAA 4268
1441 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGGTGGAAATCCCGAAGTGCAGTACACA 1500
4269 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGGTGGAAATCCCGAAGTGCAGTACACA 4328
1501 TCCAAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAACTGGACTTTTACT 1560
4329 TCCAAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAACTGGACTTTTACT 4388
1561 GAGCCTCGCCCCATTTGGCACCCTGTACCTTTACCCGTCCTCTGTA 1604
4389 GAGCCTCGCCCCATTTGGCACCCTGTACCTTTACCCGTCCTCTGTA 4432

RESULT 11
AX753251
LOCUS AX753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
ACCESSION AX753251
VERSION AX753251.1 GI:32166108
KEYWORDS Adeno-associated virus 1
SOURCE Adeno-associated virus 1
ORGANISM

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
1
REFERENCE
AUTHORS Gao, G., Wilson, J.M. and Alvira, M.
TITLE A method of detecting and/or identifying adeno-associated virus (AVV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of the University of Pennsylvania (US)
FEATURES
Location/Qualifiers
1..4718
/organism="Adeno-associated virus 1"
/mol_type="unassigned DNA"
/db_xref="taxon:85106"
ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTACGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCGCCAGGAGTGGGT 60
Db 2829 ATGGCTTACGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCGCCAGGAGTGGGT 2888
QY 61 AATGGCTCAGGAAATTTGGCATTTGCCATTGCCATTCACATGCTGGCGACAGAGTTCATCACACC 120
Db 2889 AATGGCTCAGGAAATTTGGCATTTGCCATTGCCATTCACATGCTGGCGACAGAGTTCATCACACC 2948
QY 121 AGCACCCGACCTGGGCCCTTGGCCACCTACAAATAACCACTCTCAAGCAAAATCTCCAGT 180
Db 2949 AGCACCCGACCTGGGCCCTTGGCCACCTACAAATAACCACTCTCAAGCAAAATCTCCAGT 3008
QY 181 GCTTCAACCGGGGGCGCAGCAACCAACCACTCTTCGGCTCAGCAACCCCTCTGGGGGTAT 240
Db 3009 GCTTCAACCGGGGGCGCAGCAACCAACCACTCTTCGGCTCAGCAACCCCTCTGGGGGTAT 3068
QY 241 TTTGATTTTCAACAGATTTCCACTGGCCACTTTTCCACAGCTGACTGGCAGGACTCATCAAC 300
Db 3069 TTTGATTTTCAACAGATTTCCACTGGCCACTTTTCCACAGCTGACTGGCAGGACTCATCAAC 3128
QY 301 AACAAATTTGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAAGTCAAG 360
Db 3129 AACAAATTTGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAAGTCAAG 3188
QY 361 GAGGTCAACGAAATGATGGCGTCAACCACTGCTTAATACCTTACAGCAGCGTTCAA 420
Db 3189 GAGGTCAACGAAATGATGGCGTCAACCACTGCTTAATACCTTACAGCAGCGTTCAA 3248
QY 421 GTCTTCTCGACTCGGAGTACAGACTTCCGCTCCTCGGCTCTCGGACCAAGGCTGC 480
Db 3249 GTCTTCTCGACTCGGAGTACAGACTTCCGCTCCTCGGCTCTCGGACCAAGGCTGC 3308
QY 481 CTCCCTCCGTTCCCGCGGACGTTGTTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 540
Db 3309 CTCCCTCCGTTCCCGCGGACGTTGTTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 3368
QY 541 AATGGCAGCAAGCGGTGGGACGTTTCATCTTTTACTGCTTGAATATTTCCCTTCTCAG 600
Db 3369 AATGGCAGCAAGCGGTGGGACGTTTCATCTTTTACTGCTTGAATATTTCCCTTCTCAG 3428
QY 601 ATGCTGAGAACCGGGCAACAACTTTTACTCTTCAGTACACTTTTGGAGAGTGCCTTTCCAC 660
Db 3429 ATGCTGAGAACCGGGCAACAACTTTTACTCTTCAGTACACTTTTGGAGAGTGCCTTTCCAC 3488
QY 661 AGCAGCTACGCGCACAGCCAGAGCCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 3489 AGCAGCTACGCGCACAGCCAGAGCCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 3548
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 780
Db 3549 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 3608
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGTGAACC 840
Db 3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGTGAACC 3668

QY	841	TGTTATCGGACGACGCGCTTTCTAAACAAACAAACAGACAAACAGACGAAATTTTACC	900
Db	3669	TGTTATCGGACGACGCGCTTTCTAAACAAACAAACAGACAAACAGACGAAATTTTACC	3728
QY	901	TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCATCATCAACCTGGCACT	960
Db	3729	TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCATCATCAACCTGGCACT	3788
QY	961	GCTATGGCTCTACACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTGATTT	1020
Db	3789	GCTATGGCTCTACACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTGATTT	3848
QY	1021	TTTGGAAAGAGAGCGCCGGAGCTTCAACACATGCAATGGCAATGTCATGATTAACAGAC	1080
Db	3849	TTTGGAAAGAGAGCGCCGGAGCTTCAACACATGCAATGGCAATGTCATGATTAACAGAC	3908
QY	1081	GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140
Db	3909	GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	3968
QY	1141	AATTTCCAGAGCAGCAGCAGACACCTCGCAGCGGAGATGTGATGCTATGGAGCATTA	1200
Db	3969	AATTTCCAGAGCAGCAGCAGACACCTCGCAGCGGAGATGTGATGCTATGGAGCATTA	4028
QY	1201	CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCGAGGTCCCATTTGGGCCAAATTT	1260
Db	4029	CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCGAGGTCCCATTTGGGCCAAATTT	4088
QY	1261	CCTCACACAGATGGACACTTTTACCCCGTCTCTCTTATGGCGGCTTTGGACTCAAGAAC	1320
Db	4089	CCTCACACAGATGGACACTTTTACCCCGTCTCTCTTATGGCGGCTTTGGACTCAAGAAC	4148
QY	1321	CGGCTCTCTCAGATCCTCATCAAAACACGCGCTTTCTCGGAATCCTCCGGCGGAGTTT	1380
Db	4149	CGGCTCTCTCAGATCCTCATCAAAACACGCGCTTTCTCGGAATCCTCCGGCGGAGTTT	4208
QY	1381	TCAGCTACAAAGTTTGCTTTCATTCATACCCCACTACCCCACTACCCCACTACCCCACT	1440
Db	4209	TCAGCTACAAAGTTTGCTTTCATTCATACCCCACTACCCCACTACCCCACTACCCCACT	4268
QY	1441	ATTGAATGGAGCTGCAGAAAGAAACAGACGCGCTGGAATCCCGAAGTGCAGTACACA	1500
Db	4269	ATTGAATGGAGCTGCAGAAAGAAACAGACGCGCTGGAATCCCGAAGTGCAGTACACA	4328
QY	1501	TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTTTATACT	1560
Db	4329	TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTTTATACT	4388
QY	1561	GAGCCTGGCCCAATGGCACCCGTTACCTTACCCGTCCCTCTGTA	1604
Db	4389	GAGCCTGGCCCAATGGCACCCGTTACCTTACCCGTCCCTCTGTA	4432
RESULT 12			
AF063497			
LOCUS	AF063497	4718 bp	DNA linear VRL 27-APR-1999
DEFINITION	Adeno-associated virus 1, complete genome.		
ACCESSION	AF063497		
VERSION	AF063497.1	GI:4689096	
KEYWORDS			
SOURCE	Adeno-associated virus 1		
ORGANISM	Adeno-associated virus 1		
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.		
AUTHORS	Xiao, W., Chirmule, N., Berta, S.C., McCullough, B., Gao, G. and Wilson, J.M.		
TITLE	Gene therapy vectors based on adeno-associated virus type 1		
JOURNAL	J. Virol. 73 (5), 3994-4003 (1999)		
PUBMED	10196295		
REFERENCE	2 (bases 1 to 4718)		
AUTHORS	Xiao, W. and Wilson, J.M.		
TITLE	Direct Submission		

JOURNAL	Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA		
FEATURES	Location/Qualifiers		
source	1..4718		
	/organism="Adeno-associated virus 1"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:85106"		
	/note="AAV1"		
	335..2206		
CDS	/codon_start=1		
	/product="nonstructural protein"		
	/protein_id="AAD27758.1"		
	/db_xref="GI:4689098"		
	/translation="MPGYEIVIKVPSLDLDEHLPGISDSFVSVVAEKWEWLPDSDMDLNLQALQAEKLDLQVLRVSKAPALFVQEKESYFHLHILVETTVGKSVLGRFLQIRDLKLVQITIRGIEPTLPNMFATKTRNAGGNGKVVDCYIPNYLLPKTQPELQAWTMEYISACLNLAKRLVAOHLTVSQTQONKENLPSNDAPVIRSQTARYMBELVGLVDRTISBQWQIODEQASVIFSNAAASNSRQIKAAALDNAGKIMALTKSAPDXLVGPAPPADIKTNRIYRIELUNGVEPAGVSVFLGWAOKRFGKNTIWLPGPATGKTINAEIAIAHAYPPGYCVNNTNENFFPNDVCDKVIWWEQKMTAKVESAAILGSKVRVDQKSSAQIDTPVTITNTNCAVIDGNTSTTFHQQLQDQRMFKFELTRLEHDFGKVTQGEVFFRMAQDHVTEVAHEFYRKGANKRPAPDADKSEPKRACPSVADPSTDAEGPADFADRYQNKSRHAGMLQMLFPCKTCERMNQFICTFHGT		
	RDCSECFGVSESQPVVRKTRVRLCAIHLGLRAPEIACSAACDLVVDLDDCVSEQ"		
	2223..4433		
	/codon_start=1		
	/product="capsid protein"		
	/protein_id="AAD27757.1"		
	/db_xref="GI:4689097"		
	/translation="MAADGYLPDMLNDLNSIGIREWMDLKPGAPKPKANQKQDDRGRLVLPQYKVLGPNGLDKGEVNAADAALHDKAYDQOLKAGDNPYLRYNHADAEFQERLEDTSGNGLRAVFOAKRVRLPEGLVEGAKTAPGKKRPVQSQFQEPDSSSGIGKIQDPAKKRLNFGTDSVPDQPLGEPPTAAVGPITMAAGGGAGMADNNEGADGVNAGSNHCDSTMLGDRVITTTSTWALPTYNHLYKQISSASTGASNDVHTTIA		
	STPMGYDFNRFHCFSPDRQLNNNNRPKRLNFKLFNIQVKEVTTNDVHTTIA		
	NNLTSTVOVSDSEYOLPVLGSAHQGLPPPADVPMIPOGYTLANGSOAVGRSS		
	FYCLRYFPQMLRTGNFTSYTPEEVFPHSSYAHOSLDRLMNLIDQYLYLNRTO		
	NQSQAQNKDLFPKSGSPAGMSVQPKWLPQPCYRQQRVSKTKTNNNSNFTWTGASK		
	YNLNRESIINPGTAMASHKDDDEKFFPMGSMVITPKGSAGVSNALTONIMITDSEEI		
	KATNPVATREFGTVAVNFQSSDTPDGTGVMAGALPGMWDQVDVYLGQPIWAKIPH		
	TDGHFHPSPGLMGFGGLNPPQIILIKNTVPVNPAPBFSATKSFASFITQYSTGVQSV		
	IEWELQKENSKRWNPVQVYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTPRL"		
ORIGIN			
Query Match	100.0%;	Score 1604;	DB 13; Length 4718;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1604;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAAGAGGCGCGCAGGAGTGGGT	60
Db	2829	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAAGAGGCGCGCAGGAGTGGGT	2888
QY	61	AATGCTCAGGAAATTTGGCATTTCCATGCTGGCTGGGCGCAGAGTCAATCCACC	120
Db	2889	AATGCTCAGGAAATTTGGCATTTCCATGCTGGCTGGGCGCAGAGTCAATCCACC	2948
QY	121	AGCACCGCACCTGGCGCTTGCACCTACCAATAACACCTCTTACAAGCAAAATCTCACT	180
Db	2949	AGCACCGCACCTGGCGCTTGCACCTACCAATAACACCTCTTACAAGCAAAATCTCACT	3008
QY	181	GCTTCAACGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT	240
Db	3009	GCTTCAACGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT	3068
QY	241	TTTGATTTTCAACAGATTTCCATGCGCACTTTTCCACCACTGACTGGCAGCGACTCATCAAC	300
Db	3069	TTTGATTTTCAACAGATTTCCATGCGCACTTTTCCACCACTGACTGGCAGCGACTCATCAAC	3128
QY	301	AAACAATGGGGATTCGGGCCCAAGAGACTCAACTCTTCAAACTCTTCAACATCAAGTCAAG	360
Db	3129	AAACAATGGGGATTCGGGCCCAAGAGACTCAACTCTTCAAACTCTTCAACATCAAGTCAAG	3188
QY	361	GAGGTACACGAATGATGGCGCTCACACCACTCGCTTAATAACCTTTACCGACGCGTTCAA	420

Db 3189 GAGGTCAACGAAATGATGGCGTCACAAACATCGCTAATAAATCTTACCAGCAGCGTTCAA 3248
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCTCGGCTCTGCGCACCAAGGGCTGC 480
Db 3249 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCTCGGCTCTGCGCACCAAGGGCTGC 3308
QY 481 CTCCCTCCGCTTCGCGCGGACGCTGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Db 3309 CTCCCTCCGCTTCGCGCGGACGCTGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 3368
QY 541 AATGGCAGCAAGCCGCTGGAGCGTTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 600
Db 3369 AATGGCAGCAAGCCGCTGGAGCGTTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 3428
QY 601 ATGCTGAGAACGGGCAACAATCTTACCTTCAGCTACACCTTTGAGGAAGTGGCTTCCAC 660
Db 3429 ATGCTGAGAACGGGCAACAATCTTACCTTCAGCTACACCTTTGAGGAAGTGGCTTCCAC 3488
QY 661 AGCAGCTACGCGCACAGCCGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 3489 AGCAGCTACGCGCACAGCCGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 3548
QY 721 CTGTATTACTGAACAGAACTCAAAATCAAGTCGGAAGTGCCCAAAACAAAGACTTGGCTG 780
Db 3549 CTGTATTACTGAACAGAACTCAAAATCAAGTCGGAAGTGCCCAAAACAAAGACTTGGCTG 3608
QY 781 TTTAGCCGTGGTCTCCAGCTGCGATGCTGTTCAGCCCAAAACCTGGCTACTCGGACCC 840
Db 3609 TTTAGCCGTGGTCTCCAGCTGCGATGCTGTTCAGCCCAAAACCTGGCTACTCGGACCC 3668
QY 841 TGTATTCGCGCAGCAGCGCGTTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 900
Db 3669 TGTATTCGCGCAGCAGCGCGTTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 3728
QY 901 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
Db 3729 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 3788
QY 961 GCTATGGCTCACACAAAGCAGCAGAGCAAGTTCCTTCCCATGAGCGGTGCATGATT 1020
Db 3789 GCTATGGCTCACACAAAGCAGCAGAGCAAGTTCCTTCCCATGAGCGGTGCATGATT 3848
QY 1021 TTTGGAAGAGAGCCGCGAGCTTCAACACATGCAATGGACATGTCATGATTCACAGAC 1080
Db 3849 TTTGGAAGAGAGCCGCGAGCTTCAACACATGCAATGGACATGTCATGATTCACAGAC 3908
QY 1081 GAAGAGGAAATTAAAGCCTAAACCCCTGTGGCCACCCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 3909 GAAGAGGAAATTAAAGCCTAAACCCCTGTGGCCACCCGAAAGATTTGGGACCGTGGCAGTC 3968
QY 1141 AATTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGAGCATTA 1200
Db 3969 AATTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGAGCATTA 4028
QY 1201 CCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGCGAGGTCCTCAATTTGGGCCAAAATT 1260
Db 4029 CCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGCGAGGTCCTCAATTTGGGCCAAAATT 4088
QY 1261 CCTCACACAGATGGACATTTTCAACCCGCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320
Db 4089 CCTCACACAGATGGACATTTTCAACCCGCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 4148
QY 1321 CCGCCTCTCAGATCCTCATCAAAACACAGCCTGTTCCTCGGAATCCTCGGCGGAGTTT 1380
Db 4149 CCGCCTCTCAGATCCTCATCAAAACACAGCCTGTTCCTCGGAATCCTCGGCGGAGTTT 4208
QY 1381 TCAGCTACAAGTTTGGCTTCAATTCATCAACCAATCTCCACAGACAAAGTGAAGTGGAA 1440
Db 4209 TCAGCTACAAGTTTGGCTTCAATTCATCAACCAATCTCCACAGACAAAGTGAAGTGGAA 4268
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGGAATCCCGAAGTGCAGTACACA 1500

Db 4269 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCGAAAGTGCAGTACACA 4328
QY 1501 TCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATACT 1560
Db 4329 TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATACT 4388
QY 1561 GAGCCTCGCCCATTTGGACACCCGTTACCTTACCGTCCCGCTGTA 1604
Db 4389 GAGCCTCGCCCATTTGGACACCCGTTACCTTACCGTCCCGCTGTA 4432

RESULT 13
AR527492
LOCUS Sequence 11 from patent US 6723551.
DEFINITION AR527492
ACCESSION AR527492
VERSION AR527492.1 GI:53914590
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7447)
AUTHORS Kotin, R. M., Urabe, M. and Ding, C.-T.
TITLE Production of adeno-associated virus in insect cells
JOURNAL Patent: US 6723551-A 11 20-APR-2004;
The United States of America as represented by the Department of
Health and Human Services; Washington, DC

FEATURES
source
1..7447
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1604; DB 6; Length 7447;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGAGTGGGT 60
Db 5227 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGAGTGGGT 5286
QY 61 AATGCTCAGGAAATTTGGCATTTCCGATTCACATGGCTGGCGCAGAGTCAACACC 120
Db 5287 AATGCTCAGGAAATTTGGCATTTCCGATTCACATGGCTGGCGCAGAGTCAACACC 5346
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACAATAACCACTCTACAAGCAAAATCTCCAGT 180
Db 5347 AGCACCCGACCTGGGCGCTTGGCCACCTACAATAACCACTCTACAAGCAAAATCTCCAGT 5406
QY 181 GCTTCAACCGGGGCCAGCAACCACTACTTCCGGCTACAGCAACCCCTGGGGGTAT 240
Db 5407 GCTTCAACCGGGGCCAGCAACCACTACTTCCGGCTACAGCAACCCCTGGGGGTAT 5466
QY 241 TTTGATTTCAACAGATTCACATGCGCATTTTACCAAGTGAATGGGAGAGCTCATCAAC 300
Db 5467 TTTGATTTCAACAGATTCACATGCGCATTTTACCAAGTGAATGGGAGAGCTCATCAAC 5526
QY 301 AACAAATGGGATTCGCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360
Db 5527 AACAAATGGGATTCGCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 5586
QY 361 GAGGTCAACGCAATGATGGCGTCAACCACTCGCTTAATACCTTACAGCAGCGTTCAA 420
Db 5587 GAGGTCAACGCAATGATGGCGTCAACCACTCGCTTAATACCTTACAGCAGCGTTCAA 5646
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGCTCGTCAAGTCTCGGCTCTGCGCACCAAGGGCTGC 480
Db 5647 GTCTTCTCGGACTCGGAGTACAGCTTCGCTCGTCAAGTCTCGGCTCTGCGCACCAAGGGCTGC 5706
QY 481 CTCCTCTCGTTCGCGGGGAGCGTGTTCATGATTCGCAATACCGCTACCTGACGCTCAAC 540
Db 5707 CTCCTCTCGTTCGCGGGGAGCGTGTTCATGATTCGCAATACCGCTACCTGACGCTCAAC 5766

QY 541 AATGCGAGCAAGCGGTGGAGCTTATCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 600
Db |||||||
QY 5767 AATGCGAGCAAGCGGTGGAGCTTATCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 5826
Db |||||||
QY 601 ATGCTGAGAACGGGCAACAATTTTACCTTACGCTACACCTTTTGAGGAAGTGCCTTTCCAC 660
Db |||||||
QY 5827 ATGCTGAGAACGGGCAACAATTTTACCTTACGCTACACCTTTTGAGGAAGTGCCTTTCCAC 5886
QY 661 AGCAGTACGCGACAGCGAGCGCTGGACCGGCTGATGAATCTCTCATACCAATAC 720
Db |||||||
QY 5887 AGCAGTACGCGACAGCGAGCGCTGGACCGGCTGATGAATCTCTCATACCAATAC 5946
QY 721 CTGTATTACCTCAACAGAACTCAAATCAGTCCGGAAGTGCCCAAAACAGGACTTGCTG 780
Db |||||||
QY 5947 CTGTATTACCTGAACAGAACTCAAATCAGTCCGGAAGTGCCCAAAACAGGACTTGCTG 6006
QY 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGGCCAAAACCTGGCTACTCGACCC 840
Db |||||||
QY 6007 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGGCCAAAACCTGGCTACTCGACCC 6066
QY 841 TGTATTGCGCAGAGCGGCTTTCTAAACAAAACAGACAAACAAACAGCAATTTTACC 900
Db |||||||
QY 6067 TGTATTGCGCAGAGCGGCTTTCTAAACAAAACAGACAAACAAACAGCAATTTTACC 6126
QY 901 TGGACTGGTGTCTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960
Db |||||||
QY 6127 TGGACTGGTGTCTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 6186
QY 961 GCTATTGGCTCTACACAAAGACAGCAAGTCTTTTCCATGAGCGGTGTCTATGATT 1020
Db |||||||
QY 6187 GCTATTGGCTCTACACAAAGACAGCAAGTCTTTTCCATGAGCGGTGTCTATGATT 6246
QY 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGTGANTGGACAAATGTCTATGATTAACAGAC 1080
Db |||||||
QY 6247 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGTGANTGGACAAATGTCTATGATTAACAGAC 6306
QY 1081 GAAGAGGAATTAAGCCCTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
Db |||||||
QY 6307 GAAGAGGAATTAAGCCCTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 6366
QY 1141 AATTTTCAGAGCAGCAGACAGACCTTGGGACCGGAGATGTGATCTATGGGAGCATTA 1200
Db |||||||
QY 6367 AATTTTCAGAGCAGCAGCAGACCTTGGGACCGGAGATGTGATCTATGGGAGCATTA 6426
QY 1201 CTTGGCAATGGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCAAAAT 1260
Db |||||||
QY 6427 CTTGGCAATGGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCAAAAT 6486
QY 1261 CCTCACAGATGGACACTTTCACCGCTCTCTCTTATGGGCGGCTTGGACTCAAGAAC 1320
Db |||||||
QY 6487 CCTCACAGATGGACACTTTCACCGCTCTCTCTTATGGGCGGCTTGGACTCAAGAAC 6546
QY 1321 CGGCTCTCAGATCTCATCAAAAACAGCGCTGTTCCTGCGAATCTCTCGGCGGAGTTT 1380
Db |||||||
QY 6547 CGGCTCTCAGATCTCATCAAAAACAGCGCTGTTCCTGCGAATCTCTCGGCGGAGTTT 6606
QY 1381 TCAGCTACAAAGTTGCTTCTATTCATCAACCCAAATCTCCACAGGACAAAGTGTGGAA 1440
Db |||||||
QY 6607 TCAGCTACAAAGTTGCTTCTATTCATTCACCCAAATCTCCACAGGACAAAGTGTGGAA 6666
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAGCGCTGGAAATCCCGAAGTGCAGTACACA 1500
Db |||||||
QY 6667 ATTGAATGGAGCTGCAGAAAGAAACAGCAGCGCTGGAAATCCCGAAGTGCAGTACACA 6726
QY 1501 TCCAAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAACATGGACTTTTACT 1560
Db |||||||
QY 6727 TCCAAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAACATGGACTTTTACT 6786
QY 1561 GAGCCTCGCCCAATTTGGCAACCGTTTACCTTACCCGTCCTCTGTA 1604
Db |||||||
QY 6787 GAGCCTCGCCCAATTTGGCAACCGTTTACCTTACCCGTCCTCTGTA 6830

RESULT 14
CS073493
LOCUS CS073493 2211 bp DNA linear PAT 05-MAY-2005
DEFINITION Sequence 81 from Patent WO2005033321.
ACCESSION CS073493
VERSION CS073493.1 GI:63090472
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1
AUTHORS Wilson, J.M., Gao, G., Alvita, M.R. and Vandenbergh, L.H.
TITLE Adeno-associated virus (aav) clades, sequences, vectors containing same, and uses therefor
JOURNAL Patent: WO 2005033321-A 81 14-APR-2005;
The Trustees of the University of Pennsylvania (US)
FEATURES
Location/Qualifiers
1..2211
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="new AAV serotype, clone hu.44"
ORIGIN
Query Match 99.8%; Score 1600.8; DB 6; Length 2211;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1602; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 60
Db |||||||
QY 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 666
QY 61 AATGCTCAGGAAATTTGGCAATTCGATTCACATGGCTGGGCGACAGAGTCTACACC 120
Db |||||||
QY 667 AATGCTCAGGAAATTTGGCAATTCGATTCACATGGCTGGGCGACAGAGTCTACACC 726
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACCAATACCACTCTACAAGCAATCTCCAGT 180
Db |||||||
QY 727 AGCACCCGACCTGGGCGCTTGGCCACCTACCAATACCACTCTACAAGCAATCTCCAGT 786
QY 181 GCTTCAACGGGGCGCGACCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 240
Db |||||||
QY 787 GCTTCAACGGGGCGCGACCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 846
QY 241 TTTGATTTCAACAGATTTCACTGGCCACTTTTCCACAGCTGTGCGAGGACTCATCAAC 300
Db |||||||
QY 847 TTTGATTTCAACAGATTTCACTGGCCACTTTTCCACAGCTGTGCGAGGACTCATCAAC 906
QY 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACCTTTCAAATCCCAAGTCAAG 360
Db |||||||
QY 907 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACCTTTCAAATCCCAAGTCAAG 966
QY 361 GAGGTCAAGCAAGATGATGGCGTCAACACCATCGCTAATAATCTTACCAGCAGGTTCAA 420
Db |||||||
QY 967 GAGGTCAAGCAAGATGATGGCGTCAACACCATCGCTAATAATCTTACCAGCAGGTTCAA 1026
QY 421 GTCTCTCGGAATCGGAGTACAGCTTCGTAAGTCTCGGCTCTGCGGACACGAGGCTGC 480
Db |||||||
QY 1027 GTCTCTCGGAATCGGAGTACAGCTTCGTAAGTCTCGGCTCTGCGGACACGAGGCTGC 1086
QY 481 CTCCCTCTCGGTTCCCGCGGACCTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540
Db |||||||
QY 1087 CTCCCTCTCGGTTCCCGCGGACCTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 1146
QY 541 AATGGCAGCAACCGGTGGGACGTTTATCTTTTACTGCTTGGAAATTTTCCCTTCTCAG 600
Db |||||||
QY 1147 AATGGCAGCAACCGGTGGGACGTTTATCTTTTACTGCTTGGAAATTTTCCCTTCTCAG 1206
QY 601 ATGCTGAGAACGGGCAACAACTTTTACTTACCTTACCTTACCTTTGAGGAAGTGCCTTTCCAC 660
Db |||||||
QY 1207 ATGCTGAGAACGGGCAACAACTTTTACTTACCTTACCTTACCTTTGAGGAAGTGCCTTTCCAC 1266
QY 661 AGCAGCTACGCGCACAGCCAGGCGCTGGAGCGGCTGATGAATCTCTCATCGACCAATAC 720

Db 1267 AGCAGCTACGCGCACAGCGAGCTGGACCGCGTGATGAATCCTCTCATCGACCAATAC 1326
QY 721 CTGTATTACCTGACACAGACTCAAAATCAGTCGGGAAGTGCCTCAAAACAGGACTTGGCTG 780
Db 1327 CTGTATTACCCGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAGGACTTGGCTG 1386
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACCTGGCTACTCGGACCC 840
Db 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACCTGGCTACTCGGACCC 1446
QY 841 TGTATTCCGAGCAGCGCGTCTTAAACAAAAAAGACAAACAAACAGCAATTTTACC 900
Db 1447 TGTATTCCGAGCAGCGCGTCTTAAACAAAAAAGACAAACAAACAGCAATTTTACC 1506
QY 901 TGGACTGGTCTTCAAAATATACCTCAATGGCGGTAATCCATCATCAACCTCGGCACT 960
Db 1507 TGGACTGGTCTTCAAAATATACCTCAATGGCGGTAATCCATCATCAACCTCGGCACT 1566
QY 961 GCTATGGCTTCACACAAAGACGAGCAAGTCTTCTTCCATGAGCGGTGTCATGATT 1020
Db 1567 GCTATGGCTTCACACAAAGACGAGCAAGTCTTCTTCCATGAGCGGTGTCATGATT 1626
QY 1021 TTTGAAAAAGAGCGCCGAGCTTCAAAACACTGCAATGGACAATGTCAATTAACAGAC 1080
Db 1627 TTTGAAAAAGAGCGCCGAGCTTCAAAACACTGCAATGGACAATGTCAATTAACAGAC 1686
QY 1081 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140
Db 1687 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1746
QY 1141 AATTTCCAGAGCAGCAGCAGACCTCGACCGGAGATGTGTCATCTATGGGAGCATTA 1200
Db 1747 AATTTCCAGAGCAGCAGCAGACCTCGACCGGAGATGTGTCATCTATGGGAGCATTA 1806
QY 1201 CCTGGCATGGTGTGGCAAGTAGAGACGTGTACTCGAGGTCCTCATTTGGGCCAAAAAT 1260
Db 1807 CCTGGCATGGTGTGGCAAGTAGAGACGTGTACTCGAGGTCCTCATTTGGGCCAAAAAT 1866
QY 1261 CCTCACAGATGGACACTTTCACCCGTCTCTTATGGCGGCTTTGGACTCAGAAC 1320
Db 1867 CCTCACAGATGGACACTTTCACCCGTCTCTTATGGCGGCTTTGGACTCAGAAC 1926
QY 1321 CGCCCTCTCAGATCTCATCAAAACACGCTGTCTCTCGCAATCTCTCGCGGAGTTT 1380
Db 1927 CGCCCTCTCAGATCTCTCATCAAAACACGCTGTCTCTCGCAATCTCTCGCGGAGTTT 1986
QY 1381 TCAGCTACAAAGTTTCTTCAATCATCAACCAATCTCCACAGGACAAAGTGTGGAA 1440
Db 1987 TCAGCTACAAAGTTTCTTCAATCATCAACCAATCTCCACAGGACAAAGTGTGGAA 2046
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 1500
Db 2047 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 2106
QY 1501 TCCAATTTGCAAAATCTGCAACGTTGATTTTACTGTGACAAACAAATGGACTTTTACT 1560
Db 2107 TCCAATTTGCAAAATCTGCAACGTTGATTTTACTGTGACAAACAAATGGACTTTTACT 2166
QY 1561 GAGCCTCGCCCATTTGGCACCCGTTACTTTACCCGCTCCCTGTGA 1604
Db 2167 GAGCCTCGCCCATTTGGCACCCGTTACTTTACCCGCTCCCTGTGA 2210

RESULT 15
LOCUS AY530607 2211 bp DNA linear VRL 24-JUN-2004
DEFINITION Adeno-associated virus isolate hu.44 capsid protein VP1 (cap) gene, complete cds.
ACCESSION AY530607
VERSION AY530607.1 GI:46487860
KEYWORDS Adeno-associated virus

ORGANISM Adeno-associated virus
REFERENCE 1 (bases 1 to 2211)
AUTHORS Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X. and Wilson, J.M.
TITLE Clades of Adeno-associated viruses are widely disseminated in human tissues
JOURNAL J. Virol. 78 (12), 6381-6388 (2004)
PUBMED 15163731
REFERENCE 2 (bases 1 to 2211)
AUTHORS Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X. and Wilson, J.M.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-2004) Gene Therapy Program, Division of Medical Genetics, Department of Medicine, University of Pennsylvania, 3601 Spruce Street, Wistar Institute, Philadelphia, PA 19104, USA
FEATURES
source
1..2211
/organism="Adeno-associated virus"
/mol_type="genomic DNA"
/isolate="hu.44"
/isolation_source="human tissue"
/db_xref="taxon:272636"
/country="USA"
/note="Clade A"
1..2211
/gene="cap"
1..2211
/gene="cap"
/codon_start=1
/product="capsid protein VP1"
/protein_id="AAS9292.1"
/translation="MAADGYPDWLEDTLSEGIROWWKLPGPPPPKPAERHKDSRG
LVLPGYKILGPNGLDKGEVNEADAALHDKAYDROLDSDGNVLYKYNHADAERQIG
RKEDTSTGGNIGRAVFAQKRVLPGLIVERGAETAPGKRPVQSGPQGPSSSGIG
KTQQPAKKRLNFGQTSVDPDPLGEPATPAAVGPTTMASSGGGAPMADNNSGA
DVGASGNHWCDSWLDGRVITTRTWALPTNNHLYKQISSASTGNNDHNYFY
STPMGYFDNPRFCHFSPRDMORLLNNNNFRKLNFKLNIQVKEVTNDGVTITIA
NNLTSTVQFSDSEYQLPVLGSAHQGLPPFPADVMIPOYGYLTNNNGSOAVGRSS
FYCLTFYSQMLRTGNNFTSYTFEVPFPHSYAHQSILDRMLNPLIDOLYLYPNRTO
NQSQAQNKDLFLSRGSPAGMSVQPNWLPGPCYQQRVSKTKTDNNNNFTWTGASK
YNLNGRESIINFGTAMASHKDDKFPFMSGVMIFGESAGASNTALDNVMTTDESEI
KATNPVATFQTVAFNSQSDTADGVHAMGALPGMVWQGRDVLQGPWTAKTIPH
TDGHFPPLMGGFGKXNPPQILIKNTVPANPAEFAFATKFPASFIQYSTQVQSV
IEMELQKENSKEWNPVQYTSYAKSANVDFTVNNGLYTEPRPIGTRYLTRPL"

Query Match 99.8%; Score 1600.8; DB 13; Length 2211;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1602; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGCTTACGGCGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 60
Db 607 ATGGCTTACGGCGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 666
QY 61 AATGCTCAGGAAATTTGGCATTCACATGCTGGCGGACAGAGTCAACCAACC 120
Db 667 AATGCTCAGGAAATTTGGCATTCACATGCTGGCGGACAGAGTCAACCAACC 726
QY 121 AGCACCCGACCTGGGCGCTTGGCCACTACCAATAACCACTCTTACAAGCAAAATCTCCAGT 180
Db 727 AGCACCCGACCTGGGCGCTTGGCCACTACCAATAACCACTCTTACAAGCAAAATCTCCAGT 786
QY 181 GCTTCAACGGGGCCAGCAACGACAAACCACTACTTCGGCTACAGCAACCCCTGGGGGTAT 240
Db 787 GCTTCAACGGGGCCAGCAACGACAAACCACTACTTCGGCTACAGCAACCCCTGGGGGTAT 846
QY 241 TTTGATTTCAACAGATTTCCACTGCCACTTTTCCACACGTGACTGGCAGGACATCATCAAC 300
Db 847 TTTGATTTCAACAGATTTCCACTGCCACTTTTCCACACGTGACTGGCAGGACATCATCAAC 906
QY 301 AACAAATGGGGATTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCCAAGTCAAG 360

Db 907 AACAAATGGGGATTTCGGGCCAAGAGAGCTCAAACTCTTCAACATCCNAAGTCAAG 966
Qy 361 GAGGTACAGCAAGAAATGATGGCGTCAACCATCGCTAATAACCTTTACCGACACGGTTCAA 420
Db 967 GAGGTACAGCAAGAAATGATGGCGTCAACCATCGCTAATAACCTTTACCGACACGGTTCAA 1026
Qy 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCCTCGGCTCTGCGCACAGGGCTGC 480
Db 1027 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCCTCGGCTCTGCGCACAGGGCTGC 1086
Qy 481 CTCCTCTCGTTCCTCGGCGAGCGTGTTCATGATTCGCGAATACGGCTACCTGACGCTCAAC 540
Db 1087 CTCCTCTCGTTCCTCGGCGAGCGTGTTCATGATTCGCGAATACGGCTACCTGACGCTCAAC 1146
Qy 541 AATGCGAGCAAGCCGTGGGACGTTTCATCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 600
Db 1147 AATGCGAGCAAGCCGTGGGACGTTTCATCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 1206
Qy 601 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTCCAC 660
Db 1207 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTCCAC 1266
Qy 661 AGCAGCTACGCGCACAGCGAGCGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
Db 1267 AGCAGCTACGCGCACAGCGAGCGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 1326
Qy 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTTGGAAAGTGTGCTG 780
Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTTGGAAAGTGTGCTG 1386
Qy 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTAGCCGCCAAAACCTGGCTACTGGACCC 840
Db 1387 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTAGCCGCCAAAACCTGGCTACTGGACCC 1446
Qy 841 TGTATTCCGCGAGCGCGCTTCTAAAAAAGAGACAAAGAGCAAGCAAGCAAGCAAGCAAGCA 900
Db 1447 TGTATTCCGCGAGCGCGCTTCTAAAAAAGAGACAAAGAGCAAGCAAGCAAGCAAGCAAGCA 1506
Qy 901 TGGACTGGTGTCTTAAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
Db 1507 TGGACTGGTGTCTTAAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 1566
Qy 961 GCTATGGCCCTCACAAAGACGACGAGACAGATGTTCTTCCCATGAGCGGTGTCTATGATT 1020
Db 1567 GCTATGGCCCTCACAAAGACGACGAGACAGATGTTCTTCCCATGAGCGGTGTCTATGATT 1626
Qy 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGCAATGGCAATGGCAATGGCAAT 1080
Db 1627 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGCAATGGCAATGGCAATGGCAAT 1686
Qy 1081 GAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGGTGGCAGTC 1140
Db 1687 GAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGGTGGCAGTC 1746
Qy 1141 AATTTCCAGAGCAGCAGCAGACCCCTGGACCGGAGATGTGCATGCTATGGAGCATTTA 1200
Db 1747 AATTTCCAGAGCAGCAGCAGACCCCTGGACCGGAGATGTGCATGCTATGGAGCATTTA 1806
Qy 1201 CCTGGCATGTGTGGCAAGATAGAGAGGTGTACCTGCGGGTCCCATTTGGGCCAAAATTT 1260
Db 1807 CCTGGCATGTGTGGCAAGATAGAGAGGTGTACCTGCGGGTCCCATTTGGGCCAAAATTT 1866
Qy 1261 CCTCACAGATGGACACTTTTCAACCGTCTCTCTTTATGGGGGCTTTGGACTCAAGAAC 1320
Db 1867 CCTCACAGATGGACACTTTTCAACCGTCTCTCTTTATGGGGGCTTTGGACTCAAGAAC 1926
Qy 1321 CGGCTCTCTCAGATCTCTCATCAAAAACAGCCCTGTTCTTGGCGAATCCTCGGCGGAGTTT 1380
Db 1927 CGGCTCTCTCAGATCTCTCATCAAAAACAGCCCTGTTCTTGGCGAATCCTCGGCGGAGTTT 1986
Qy 1381 TCAGCTACAAAGTTTCTCTTCAATCATCCCAATACCTCCACAGCAAGTGAAGTGGAA 1440

Db 1987 TCAGCTACAAAGTTTCTCTTCAATCATCCCAATACTCCACAGGACAGTGAAGTGGAA 2046
Qy 1441 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAAATCCGAAAGTGCAGTACACA 1500
Db 2047 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAAATCCGAAAGTGCAGTACACA 2106
Qy 1501 TCCAAATTATGCAAAATCTGCCAAGCTTTGATTTTACTGTGGACAACAATGGACTTTTATCT 1560
Db 2107 TCCAAATTATGCAAAATCTGCCAAGCTTTGATTTTACTGTGGACAACAATGGACTTTTATCT 2166
Qy 1561 GAGCCTCGCCCAATTTGGCACCCGTTACCTTACCCGTCGCCCTGTA 1604
Db 2167 GAGCCTCGCCCAATTTGGCACCCGTTACCTTACCCGTCGCCCTGTA 2210

Search completed: November 29, 2005, 00:05:36

Job time : 5865.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 7013.94 Seconds
(without alignments)
14735.325 Million cell updates/sec

Title: US-10-696-900-1_COPY_2223_4431
Perfect score: 2209
Sequence: 1 atggctgcgatgttatct.....ttacctaccgtccctcgt 2209

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	180.8	8.2	264	7	CO892248 BovGen_20
C 2	125.8	5.7	195	7	CO888893 BovGen_17
C 3	76.2	3.4	753	9	BH115587 RPCI-24-3
C 4	50.6	2.3	803	9	BZ265284 CH230-375
C 5	48	2.2	367	3	BJ432209 BJ432209
C 6	47	2.1	581	5	BW335572 BW335572
C 7	45.6	2.1	681	10	AL133930 Tetraodon
C 8	45.6	2.1	869	7	CK159167 FGAS04056
C 9	45.4	2.1	581	6	CF569196 EST057 Su
C 10	45.4	2.1	983	11	CN304QNN
C 11	45.2	2.0	1075	10	CW931784
C 12	45	2.0	289	1	AI736567
C 13	45	2.0	322	1	AW206887 UI-H-Bil-
C 14	45	2.0	356	2	BG170384 602322706
C 15	45	2.0	451	1	AI439550 tc90C03.x
C 16	45	2.0	454	1	AI240019 gh33h11.x
C 17	45	2.0	472	1	AA806398 oc27a06.s
C 18	45	2.0	479	1	AI566324 tq71b07.x
C 19	45	2.0	486	3	BW352611 ig68h10.x
C 20	45	2.0	488	1	AW181950 xj69a04.x
C 21	45	2.0	498	2	BF194970 7091b09.x
C 22	45	2.0	504	1	AI989403 wt80g09.x

C 23	45	2.0	511	1	AI018231
C 24	45	2.0	535	2	BF732474
C 25	45	2.0	537	1	AI761007
C 26	45	2.0	551	1	AA421597
C 27	45	2.0	612	1	AI523558
C 28	45	2.0	620	5	BX349872
C 29	45	2.0	664	2	BF058781
C 30	45	2.0	665	2	BE903379
C 31	45	2.0	762	1	AI373205
C 32	45	2.0	910	10	CNS0060N
C 33	45	2.0	1362	4	BC020427
C 34	44.8	2.0	199	1	AA864674
C 35	44.8	2.0	556	3	BI948203
C 36	44.8	2.0	821	3	BI948601
C 37	44.8	2.0	830	8	CV764236
C 38	44.8	2.0	1101	8	DR736539
C 39	44	2.0	362	1	AI870396
C 40	44	2.0	381	1	AI240354
C 41	44	2.0	579	1	AI762677
C 42	44	2.0	622	6	CB870960
C 43	44	2.0	925	10	CNS0091P
C 44	43.8	2.0	512	5	CA024678
C 45	43.8	2.0	596	1	AV601091

ALIGNMENTS

RESULT 1
CO892248/c

LOCUS
DEFINITION

ACCESSION
VERSION

KEYWORDS
SOURCE

ORGANISM
REFERENCE

AUTHORS
TITLE

JOURNAL
COMMENT

CO892248 264 bp mRNA linear EST 01-SEP-2004
BovGen.20573 normal cattle brain Bos taurus cDNA clone
RZPDp1056M0360Q 5', mRNA sequence.
CO892248
CO892248.1 GI:51922548
EST.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 264)
Hennig S., Janitz M., Herwig R. and Williams J.
Generation, annotation, evolutionary analysis and database
integration of 14969 cattle EST clusters
Unpublished (2004)
Contact: Hennig S
Laboratory 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennigmolgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONFP cluster was selected for sequencing. cDNA clones and
filters are distributed via Deutsches Ressourcenzentrum fuer
Genomforschung GmbH (<http://www.rzpd.de>).
PCR Primers
FORWARD: 5' CCCAGCTTTACACTTTATGCTTCGGGTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATAGCCAGCTGGGAAAGGGGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5'-CCGTCGCGAATCCCGGT-3' (M13RSP).

FEATURES
source

Location/Qualifiers
1..264
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="RZPDp1056M0360Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"

/clone lib="normal cattle brain"
 /note="Organ: brain; Vector: pSport1; Site 1: NotI;
 Site 2: SalI; Random primed and directionally cloned in
 pSport1 vector using NotI
 (5'-pGACTAGTTCTAGATCGGAGCGCGCC (T)15-3' and SalI 5'-
 TCGACCCACGGCTCG-3' adapters (Gibco BRL))"

ORIGIN

Query Match 8.2%; Score 180.8; DB 7; Length 264;
 Best Local Similarity 80.3%; Pred. No. 1.2e-39;
 Matches 212; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1847 GTCCATTGGGCCAAAATTCCTCACAGATGGACATTTCCACCGTCTCTCTATTGG 1906
 |||||
 Db 264 GTCCATTGGGCAATGATTCACACTGGAGCGACATTTTCACCCCTCTCCCTCATGG 205
 |||||

QY 1907 GCGGCTTTGGACTCAAGAACCGGCTCTCTCAGATCCTCATCAAAACACGCGCTGTTCCTG 1966
 |||||
 Db 204 GGGGATTGGGACTTAATCACCTCTCTCCACAGATTTTCATCAAGAACACGCGGTTCTGTG 145
 |||||

QY 1967 CQAATCTCCGGCGGAGTTTCAGCTACAAAGTTTGCTTCATTCATCACCAGTAATCTCA 2026
 |||||
 Db 144 CGAGTCTTCGAGCACGTTTCAGTGGCGCAAGTTTGCTTCATTCATCACAGTACTCCA 85
 |||||

QY 2027 CAGGACAAGTGTGTTGAAATTTGAATGGAGCTGCAGAAAGAAACACAGCAGCGCTGGA 2086
 |||||
 Db 84 CGGTCAGTGTAGCTGGAGATTGATGGAGCTGGGAGAGAAACACAGCAACGCTGGA 25
 |||||

QY 2087 ATCCGGAAGTGCAGTACATCCA 2110
 |||||
 Db 24 ATCCGGAATTCAGTACACTCCA 1
 |||||

RESULT 2

CO888893/c
 LOCUS 195 bp mRNA linear EST 01-SEP-2004
 DEFINITION BovGen.17218 normal cattle brain Bos taurus cDNA clone
 RZPDp105600960Q 5', mRNA sequence.

ACCESSION CO888893
 VERSION EST.
 KEYWORDS CO888893.1 GI:51819178
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.
 1 (bases 1 to 195)
 Hennig.S., Janitz,M., Herwig,R. and Williams,J.
 Generation, annotation, evolutionary analysis and database
 integration of 14969 cattle EST clusters
 Unpublished (2004)
 JOURNAL Contact: Hennig S
 Laboratory 123, dept.Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr.63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1612
 Fax: +49 30 8413 1380
 Email: hennigmolgen.mpg.de

COMMENT The library was characterised by oligonucleotide fingerprinting
 (ONFP) to reduce sequencing redundancy. According to the ONFP
 procedure, clones that display the same hybridisation matrix with a
 battery of 200 8mer oligonucleotides are grouped into clusters. One
 clone per ONFP cluster was selected for sequencing. cDNA clones and
 filters are distributed via Deutsches Ressourcenzentrum fuer
 Genomforschung GmbH (<http://www.rzpd.de>).

PCR Primers
 FORWARD: 5' CCCCAGGCTTACACTTTATGTTCCGGCTCG 3' (M13RSP) 5'-seq
 BACKWARD: 5' GCTATTACCGCAGCTGGGAGAGGGGATGG 3' (M13FSP) 3'-seq
 Seq primer: 5'-CCGGTCCGGAATTCCTCCGGT-3' (M13RSP).
 Location/Qualifiers
 1. .195
 /organism="Bos taurus"
 /mol_type="mRNA"

FEATURES
source

/db_xref="taxon:9913"
 /clone="RZPDp105600960Q"
 /sex="female"
 /tissue_type="brain tissue"
 /dev_stage="adult brain"
 /clone_lib="normal cattle brain"
 /note="Organ: brain; Vector: pSport1; Site 1: NotI;
 Site 2: SalI; Random primed and directionally cloned in
 pSport1 vector using NotI
 (5'-pGACTAGTTCTAGATCGGAGCGCGCC (T)15-3' and SalI 5'-
 TCGACCCACGGCTCG-3' adapters (Gibco BRL))"

ORIGIN

Query Match 5.7%; Score 125.8; DB 7; Length 195;
 Best Local Similarity 78.2%; Pred. No. 4.9e-24;
 Matches 151; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2017 CAATACTCCAGACAGTGTGGAATTCGAATGGGAGCTGCAGAAAGAAAACAGC 2076
 |||||
 Db 194 CAGTACTCCAGGACAGGTCAGCGTGGAGATCGATGGGAGTTGCAGAAGGTAACAGC 135
 |||||

QY 2077 AAGCGCTGGAATCCGAAGTGCAGTACATCCAAATTAATGCAAAATCTGCCAAGTTGAT 2136
 |||||
 Db 134 AAACGGTGGAAATCCGAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGAT 75
 |||||

QY 2137 TTTACTGTGGACAACATGGACTTTTATCTAGCTCGCCCATTTGGCACCCTTACCTT 2196
 |||||
 Db 74 TTTAATGTGGACATTAATGCGGTGATTTCAGAGCTCGCCCATTTGGCACCAGATACCTG 15
 |||||

QY 2197 ACCCGTCCCTGT 2209
 |||||
 Db 14 ACTCGTAATGTGT 2

RESULT 3

BH115587
 LOCUS 753 bp DNA linear GSS 19-JUL-2001
 DEFINITION RPCI-24-358F16.TV RPCI-24 Mus musculus genomic clone
 RPCI-24-358F16, genomic survey sequence.

ACCESSION BH115587
 VERSION GSS.
 KEYWORDS BH115587.1 GI:14954954
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 753)
 Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
 Tsagaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,
 Russell,D., de Jong,P. and Fraser,C.M.
 Mouse BAC End Sequences from Library RPCI-24
 Unpublished (1999)
 JOURNAL Other GSSs: RPCI-24-358F16.TJ
 COMMENT Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC
 library availability, please contact Pieter de Jong
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC
 Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end
 page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 358 row: F column: 16
 Seq primer: T7
 Class: BAC ends.
 Location/Qualifiers
 1. .753
 /organism="Mus musculus"
 /mol_type="genomic DNA"

FEATURES
source

```
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-350F16"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/notes="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; The
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

ORIGIN
Query Match      3.4%; Score 76.2; DB 9; Length 753;
Best Local Similarity 51.6%; Pred. No. 1e-09;
Matches 268; Conservative 0; Mismatches 223; Indels 28; Gaps 3;

QY 266 CGTACCTCGGTATACACCGCCGAGTTCAGGAGCGTCTGCAAGAGATACGT 325
DB 207 CTTGTCTCCATAAATACACAGAGCAGAGTTTCAGGAGAAACTCCAAATAC 266
QY 326 CTTTGGGGCACCTCGGGCGAGCAGTCTTCAGGCGCAAGAGCGGTTCTCGAACCTC 385
DB 267 TTTTTTTGGCA-----GGAACCGTGCCAATGCCAAGAAAGGGTTCTCAAACCCA 318
QY 386 TCGGTCTGTGTAGGAAGCGCTAAGACGGCTCTCGAAAGAAACGTCCTCGTAGAGCAGT 445
DB 319 TCGGCTTGGTCTAAGAGAGGATTACGATGCTCTCTGGAGAAAT----- 363
QY 446 CGGCACAAGAGCCAGACTCTCTCTCGGGCATCGGCAAGACAGCGCCAGCCCGCTAAAA 505
DB 364 ----AGATTCCAGCTTAAACTCTCTCCCCACATGAAGAGATACACTCGCTCTTTTCAGA 419
QY 506 AGAGACTCAATTTGGTCAGACTGGGCACTCAGAGTCAGTCCCGATCCACACCTCTCG 565
DB 420 GTTATGCCAAAAACAGTAGAGTGGAGAGCTGGCTCATTAGTCTATGACAAACAGATGTA 479
QY 566 GAGAACTCCAGCAACCCCGCTGCTGTGGGACCTACTACAATGGCTTCAGGGGGTGGCG 625
DB 480 GCAGTTTCCAGCAGATATCTCCATTTAGACACTTCTATAATCTCTGGAGCTGGAGTTC 539
QY 626 CACCAATGGCAGACAATAAGAGAGCGCCGACGAGTGGGTAAATGCCTCAGGAAATGGC 685
DB 540 TACTAATGGCAACAACAAC-CAGACACTGATGGAGTGAGCAATGCCATGGATGATGGC 598
QY 686 ATTGCGAATTCACATGCTGGGCGACAGAGTCATCACCAAGCAGCAGCCGACCTGGCCT 745
DB 599 ATTGCCAATTCGAAGTGTATGGTAGACTGAGTCAATTACCTAATCTCCCCAGACCTGGGTCC 658
QY 746 TGCCACCTACATAACCACTCTACAAAGCAAAATCTCCA 784
DB 659 TGCCCAACTCAACTAAACATTTCCACAACTATATGAACA 697

RESULT 4
BZ265284
LOCUS CH230-375A22.1 803 bp DNA linear GSS 15-OCT-2002
DEFINITION CH230-375A22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
BZ265284
ACCESSION BZ265284.1 GI:23978528
VERSION BZ265284.1
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 803)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)

TITLE
JOURNAL
```

Other_GSSs: CH230-375A22.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 375 row: A column: 22
Seq primer: SP6
Class: BAC ends.

Location/Qualifiers
1..803
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-375A22"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match 2.3%; Score 50.6; DB 9; Length 803;
Best Local Similarity 60.6%; Pred. No. 0.02;
Matches 83; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 276 GTATATACACCGCCGAGCGGAGTTTCAGGAGCGTCTGCAAGAGATACGTCTTTGGGG 335
DB 304 GTACAGCCCTGTGCAAGCAGAAATTCAGAGAAATCCAAACAGATGCTTTTCCAGAGG 363
QY 336 CAACCTCGGGGAGCAGTCTCCAGGCGGAGGTTCTCGAACCTCTCGGTCTGGT 395
DB 364 CAATCTCAACCAAGTACTCTTCAGGCTAAAAGAGACTTCTGATCCCTTTGACCTGGT 423
QY 396 TGAGGAAGCGCTAAGA 412
DB 424 TAAAAAGTTGTCCAGA 440

RESULT 5
BJ432209/c
LOCUS BJ432209 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION BJ432209 dictyostelium clone ddv18d02 3', mRNA sequence.
BJ432209
ACCESSION BJ432209.1 GI:19406931
VERSION BJ432209.1
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS 1 (bases 1 to 367)
Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

```

source
1. 367
/organism="Dictyostellium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/cloned="ddv18d02"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostellium discoideum cDNA library, VF"

ORIGIN
Query Match 2.2%; Score 48; DB 3; Length 367;
Best Local Similarity 49.6%; Pred. No. 0.085;
Matches 120; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 716 TCATCACCACGACCCGACCTGGGCGCTTCCGACCTTACCAATTAACCACTCTTACAAGC 775
Db 328 TCAACAACATCANCAACATCAANTACCTCTAAACCAACAACACTAGATCAACACATCAACT 269
QY 776 AAATCTCCAGTCTTCAACGGGGGCGGACGACGACCAACCACTACTTCCGCTCAGCACCC 835
Db 268 ACATCAACTACCTCTTAACCAACAACCAACATCAACTTACATCACTCACTCACTACATCA 209
QY 836 CTGGGGGTATTTGATTTTCAACAGATTCCACTTCCACTTTTCAACCACTGACTGGCAGC 895
Db 208 ACTACTTCCAAACCACTACTATCAACCAACATCACTACTTCCAAACCACTACTGGT 149
QY 896 GACTCATCAACAACAAATTTGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACA 955
Db 148 TCATCAACAACAACAACACTGGTTTCATCAACAACAACACTGGTTTCATCAACAACAACATCA 89
QY 956 TC 957
Db 88 CC 87

RESULT 6
BM335572 581 bp mRNA linear EST 27-MAY-2004
LOCUS BM335572 Yutaka Satou unpublished cDNA library, adult digestive
DEFINITION gland Clona intestinalis cDNA clone cidg851f5', mRNA sequence.
ACCESSION BM335572.1 GI:47747373
VERSION 1 (bases 1 to 581)
KEYWORDS Expressed genes in Clona intestinalis (2004)
SOURCE Contact: Yutaka Satou
ORGANISM Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4095
Fax: 81-75-705-1113
Email: yutaka@ascidian.zool.kyoto-u.ac.jp.
REFERENCE 1. 581
Location/Qualifiers
1. 581
/organism="Clona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/cloned="cidg851f5"
/tissue_type="digestive gland"
/dev_stage="adult"
/clone_lib="Yutaka Satou unpublished cDNA library, adult
digestive gland"

ORIGIN
Query Match 2.1%; Score 47; DB 5; Length 581;
Best Local Similarity 45.9%; Pred. No. 0.19;

source
1. 681
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"

ORIGIN
Query Match 2.2%; Score 48; DB 3; Length 681;
Best Local Similarity 49.6%; Pred. No. 0.085;
Matches 161; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 699 GCGATTCACATGGCTGGGCGACAGAGTCATCACACGACCCGCGACCTGGGCTTGGC 748
Db 148 GCAACTACAGAAGCACCACCAACATCATCAACAACAACAGCGGCGCAACAACAGCAACA 207
QY 749 CCACCTCAATAACCACTCTTACAGCAAACTCCAGTGCTTCAACGGGGGCGGACGACG 808
Db 208 CCAACAACCAACAACAACAACCAACCAACAGCAGCAACAACAGCAACAACATCACAACA 267
QY 809 ACAACCACTACTTGGCTTACAGCACCCCTGGGGGTATTTTGAATTTCAACAGATTCAC 868
Db 268 ACAACGATCATCAACAACAACAACAGCATCAACAACAACAGCAACAACAGCAACAACA 327
QY 869 GCCACTTTTACCAGTGTGCTGGCAGCGACTCATCAACAACAATTTGGGGATTCGGGCCA 928
Db 328 GGAACAACAACGGCACCACCAACAACAACAACAGCAACAACAACAGCAACAACAGCAACA 387
QY 929 AGAGACTCAACTTCAAACTCTTCAACATCCAGTCAAGGAGGTCAAGCAAGATGATGGG 988
Db 388 ACAACAGAAACAACAACAGCACCACCAAGAAACAACAGCAACAACAGCAACAATACATCA 447
QY 989 TCACAACCATCGCTAATAAATTAACCTTACCAGCAGGTTCAGTCTTCTCGGACT 1039
Db 448 ACAACAACAACGCTAATAACAAGAACACACAACAACAACATCACCACCACT 498

RESULT 7
CNS02EOD 681 bp DNA linear GSS 01-SEP-2000
LOCUS CNS02EOD/1 GI:7832096
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
262H14 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL193990
VERSION 1
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1. Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bertot, A., Fzames, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W., and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
2. Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bertot, A., and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
3. (bases 1 to 681)
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1. 681
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"

```



```

/clone="262H14"
/clone_lib="G"
/notes="Genoscope sequence ID : COAG362DD07LP1
end : T7"

```

ORIGIN

```

Query Match      2.1%; Score 45.6; DB 10; Length 681;
Best Local Similarity 43.4%; Pred. No. 0.51;
Matches 144; Conservative 10; Mismatches 178; Indels 0; Gaps 0;

QY 637 GACAATACGAAGCGCGCGAGTGGGTAAATCCCTCAGGAATTTGGCATTC 696
DB 351 GACAACACGACGACGACGACGACGACGACGACGACGACGACGACGAC 292
QY 697 ACATGGCTGGCGGACGACGAGTATCACCACGACGACGACGACGACGACGAC 756
DB 291 RACGGGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 232
QY 757 AATAACGACCTTACGAAGCAATCTCCAGTGTCTTCAACGGGGCCGACGACGAC 816
DB 231 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 172
QY 817 TACTTGGCTAGACGACCCCTGGGGTATTTGATTTCAACAGATTCCTCCACTTT 876
DB 171 AAAAACACACACACACACACACACACACACACACACACACACACAC 112
QY 877 TCACCGCTGACTGGCAGCGACTCATCAACAACAATTGGGGATTTCGGGCCAAGAGACTC 936
DB 111 AACAAAACACACACACACACACACACACACACACACACACACACAC 52
QY 937 AACTTCAAACCTTTCAACATCCAGTCAAGGA 968
DB 51 AASMACAACACACACACACACACACACACACACACACACACACACAC

```

RESULT 8

```

CK159167/c
LOCUS      CK159167      869 bp      mRNA      linear      EST 05-DEC-2003
DEFINITION FGAS040564 Triticum aestivum FGAS: TaLts5 Triticum aestivum cDNA,
mRNA sequence.

```

ACCESSION

```

CK159167
VERSION    CK159167.1 GI:38998053

```

KEYWORDS

```

EST.

```

SOURCE

```

Triticum aestivum (bread wheat)

```

ORGANISM

```

Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

REFERENCE

```

1 (bases 1 to 869)
Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
Gensewein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,
Penniket,C., Roach,J.L. and Sarhan,F.

```

AUTHORS

```

Functional Genomics of Abiotic Stress In Wheat and Canola Crops

```

TITLE

```

Unpublished (2003)

```

JOURNAL

```

Contact: Wm L Crosby
Bioinformatics

```

COMMENT

```

University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_este@cs.usask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [128,636].
Plate: TaLts537, row: N column: 23.

```

FEATURES

source

```

1..869
/organisms="Triticum aestivum"
/mol_type="mRNA"

```

```

/cultivar="Wheat line PI 178383"
/db_xref="taxon:4565"
/lab_host="DH5 alpha"

```

```

/clone_lib="Triticum aestivum FGAS: TaLts5"

```

```

/notes="Organ: Crown; Vector: pGEM-T; SSH (suppression
subtractive hybridization) cDNA library from genotype
Pi178383 cold hardened at 2 C for 21 days and 49 days
(equal amount of cDNA pooled together before subtraction,
tester) and subtracted against genotype Norstar cold
hardened at 2 C for 1 day (24 H) (driver). Modified Smart
cDNA (Clontech) priming and non-directional cloning"

```

ORIGIN

```

Query Match      2.1%; Score 45.6; DB 7; Length 869;
Best Local Similarity 48.1%; Pred. No. 0.56;
Matches 129; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 751 ACTTACAATTAACCACTCTTACGAAGCAATCTCCAGTCTTCAACGGGGCCGACGAC 810
DB 785 AACAAACCAACACCAACCAACGACGACGACGACGACGACGACGACGAC 726
QY 811 AACCACTACTTGGCTTACAGCACCCCTGGGGTATTTTGATTTCAACAGATTCCTACTGC 870
DB 725 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 666
QY 871 CACTTTTACCACGCTGTCAGCGGACTCATCAACAACAATTGGGGATTTCGGGCCAAG 930
DB 665 CACCACAACCAACAACAACAACAACAACAACAACAACAACAACAACAACA 606
QY 931 AGACTCAACTTCAACTCTTCAACATCCAGTCAAGGAGGTACGAGGATGATGGGTC 990
DB 605 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 546
QY 991 ACAACCATCGCTAATTAACCTTACCAGCA 1018
DB 545 AACCAACAACAACAACAACAACAACAACAACAACAACAACAACA

```

RESULT 9

```

CF569196
LOCUS      CF569196      581 bp      mRNA      linear      EST 08-SEP-2004
DEFINITION EST057 Subtracted, Clontech (cat. # K1804-1) Triticum aestivum cDNA
clone FDC57 5', mRNA sequence.

```

ACCESSION

```

CF569196
VERSION    CF569196.1 GI:51921536

```

KEYWORDS

```

EST.

```

SOURCE

```

Triticum aestivum (bread wheat)

```

ORGANISM

```

Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

```

REFERENCE

```

1 (bases 1 to 581)
Xiao,K., Bai,G.H. and Carver,B.F.

```

AUTHORS

```

Nylon Filter Arrays Reveal Differential Expression of Expressed
Sequence Tags in Wheat Roots Under Aluminum Stress

```

TITLE

```

J. Integr. Plant Biol. 47 (7), 839-848 (2005)

```

JOURNAL

```

Contact: Guihua Bai
USDA/ARS and Department of Agronomy

```

COMMENT

```

Kansas State University
Manhattan, KS 66506, USA
Email: gbai@bear.agron.ksu.edu

```

```

Seq primer: M13 Forward
High quality sequence stop: 581.

```

FEATURES

source

```

1..581
/organisms="Triticum aestivum"
/mol_type="mRNA"

```

cultivar

```

/cultivar="OK91G106"

```

db_xref

```

/db_xref="taxon:4565"

```

clone

```

/clone="FDC57"

```

tissue_type

```

/tissue_type="root"

```

clone_lib

```

/clone_lib="Subtracted, Clontech (cat. # K1804-1)"

```

notes

```

/notes="EST from wheat (Triticum aestivum, cv. OK91G106)"

```

root in response to aluminum stress"

ORIGIN	Query Match	2.1%; Score 45.4; DB 6; Length 581;
	Best Local Similarity 46.6%; Pred. No. 0.55;	
	Matches 145; Conservative 0; Mismatches 166; Indels 0; Gaps 0;	
Qy	708 CGACGAGTGCATCACACACGACCCGACCTTGGGGCTTGGCCACCTACATAATACCACT 767	
Db	152 CACACAAACAAACAGACCCACCAACACGAGCAACAGCNAACAAACAGCAACACAA 211	
Qy	768 CTACAAGCAAAATCTCCAGTGTCTTCAACGGGGCCAGCAACAGCAACCACTACTTTCGGCTA 827	
Db	212 CAACAAACAAACAGCAACAAACAGCAACAAACAAACAAACATCAGCAACACAA 271	
Qy	828 CAGCAACCCCTGGGGGTATTTTGATTTCAACAGAGATTCACCTGCCACTTTTTCACCACTGA 887	
Db	272 CAGCAGACCAACACGACGACCAACACAGCAGCAGCAGCAACAGCAACAGTAAACACAGCA 331	
Qy	888 CTGGCAGCGACTCATCAACAAACAATTTGGGGATTCGGGCCCAAGAGATCTCAACTTCAAACT 947	
Db	332 CAACAGCAGCAACACCAACAAACAAACAAACAAACAAACAAACAAACAAACAA 391	
Qy	948 CTTCAACATTCOAAGTCAAGGAGGTTCACGACGAATGTATGGGTCTCAACACCATCGGTATATA 1007	
Db	392 CAGCAACAAACAAACAAACATCATCAGCAACAAACATCAGCAACAAACAGCAGCAGCA 451	
Qy	1008 CCTTACCAGCA 1018	
Db	452 CAGCAACAGCA 462	

CNS04QNN	983 bp	DNA	linear	GSS 01-SEP-2000
LOCUS				
DEFINITION	Tetraodon nigroviridis genome survey sequence T7 end of clone 193H24 of library G from Tetraodon nigroviridis, genomic survey sequence.			

sequence.
AL302828.1 GI:8183374
AL302828.1 GI:8183374
GSS; genome survey sequence.
KEYWORDS
Tetraodon nigroviridis
SOURCE
Tetraodon nigroviridis
ORGANISM
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1
REFERENCE
Rosset Colluis,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernet,A., Fzames,C., Winkler,P., Brottier,P., Quetier,P.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
10835645

2	10833843	Roest Crollius, H., Jailion, O., Dasilva, C., Ozouf-Costaz, C., Fiamas, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
TITLE		Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL		Genome Res. 10 (7), 939-949 (2000)
PUBMED		10899143
REFERENCE		3 (bases 1 to 983)
AUTHORS		GENESCOPE.

11/10/2000	Direct submission	Genoscope - Centre National de Sequencage : Submitted (12-APR-2000) FRANCE (E-mail : sege@genoscope.cns.fr BP 191 91006 EVRY cedex - FRANCE (E-mail : sege@genoscope.cns.fr - Web : www.genoscope.cns.fr)
11/10/2000	COMMENT	This sequence is a single read and was generated as part of a large scale clone and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .

FEATURES
SOURCE

```

Location/Qualifiers
1..983
/organism="Tetradodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="129H24"
/clone_lib="G"
/note="genoscope sequence ID : C0BG129DD12LP1
end : 17"
```

ORIGIN

Query Match	2.1k	Score	45.4	DB	11	Length	983
Best Local Similarity	47.1k	Pred.	No.	0.66			
Matches	139	Conservative	0	Mismatches	156	Indels	0
Gaps	0						
Qy	29	GGCTCGAGGACAACCTCTCTGAGGGATTCGCGAGTGTGGGACTTGAAACCTGGAGCCC	88				
Db	381	GCCAGGACGAGGACCAAGGCCAGGACCAAGGCCAGGACCAAGGACCAAGCAGG	322				
Qy	89	CGAAGCCCAAGCCAACCGACAAAGACGACGACCGCCGGGTCTGTGCTTCTTGCT	148				
Db	321	GCCAGGGCCAGGACCAAGGGCCAGGACTAGGGCCACGCGCCAGGACGAGCCAGG	262				
Qy	149	ACAAGTACTTCGGAACCTTCAAGCGACTCGACAAGGGGGAGCCCGTCAACGCGCGGACG	208				
Db	261	GCCAGGACCAAGGCCAGGACCAAGGGCCACGCGCCAGGCGCAGGACCAAGCAGG	202				
Qy	209	CAGGGGCCCTCGAGACGACAAGGCCCTACGACAGCGAGCTCAAGCGGGGTGACATCCGT	268				
Db	201	ACCAAGGCCAGGGCCAGGACCGGGCCAGGGCCAGGGCCAGGACCAAGGGCCAGG	142				
Qy	269	ACCTGCGGTATACCCAGCGGACCGCGAGTTTTCAGGAGCGCTCTGCAAGAAGATAC	323				
hb	141	GCCAGGCCAGGACCAAGGGCCAGGGCCAGGGCCAGGACCAAGGGCCAGGACCAAG	87				

RESULT 11
CW931784/c

LOCUS CW931784 1075 bp DNA linear GSS 24-FEB-2005
DEFINITION EDC981TF A. castellanii, 6-8 kb library from total genomic DNA
ACanthamoeba castellanii genomic clone EDC981, genomic survey
sequence.
FEATURES
Source: CW931784
Accession: CW931784.1 GI:60247281
Keywords: GSS.
Organism: Acanthamoeba castellanii
Acanthamoeba castellanii
Eukaryotes: Acanthamoebidae; Acanthamoeba.
Reference: Anderson, I.J. and Loftus, B.J.
Gene discovery in the Acanthamoeba castellanii genome
Unpublished (2004)
Contact: Iain Anderson
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-795-7949
Fax: 301-838-0208
Class: shotgun.

FEATURES

```

Location/Qualifiers
1. .1075
/organism="Acanthamoeba castellanii"
/mol_type="genomic DNA"
/strain="Neff"
/db_xref="taxon:5755"
/clone_lib="A. castellanii, 6-8 kb l
genomic DNA"
/note="Vector: pHO52"

```

ORIGIN

Query Match 2.0%; Score 45.2; DB 10; Length 1075;
Best Local Similarity 54.1%; Pred. No. 0.78;
Matches 92; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 100 GCCAACCAAGAAAGCAGACGACGGCCGGGGTCTGGTCTTCTGCTTCAAGTACCTC 159
 |||||
 Db 698 GCGACGAGGACTCGAGCTGGCCGCGCAGGTGAGCGCGCTGCAGGCCGCAAGGACCGG 639
 |||||
 QY 160 GGACCTTTCAACGGACTCGACAAAGGGGGAGCCCGTCAACGCGCGGACGACGCGGCCCTC 219
 |||||
 Db 638 CTGGCGCTCGAGTGGCCGACAGGACGAGCGCGGTCAAGGAGCTCGAGAAAGACGGCCGAC 579
 |||||
 QY 220 GAGCAGCAAGCGCTTACGACGACGCTCAAGCGGGTGACAATCCGTA 269
 |||||
 Db 578 GACCTGGTCAAGGATACGAGAGGAGAAAGCAGCGCGCGACGACCTGGA 529
 |||||

RESULT 12
 A1796567/c
 LOCUS
 DEFINITION
 wh46h04.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2383831 3' similar to TR:000613 000613 T3 RECEPTOR-ASSOCIATING COFACTOR-1. [1]
 ; mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 289)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 326 Std Error: 0.00
 Seq primer: -40UP from Gibco.

FEATURES
 source
 Location/Qualifiers
 1..289
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2383831"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Kid11"
 /notes="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid1 was prepared, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 2.0%; Score 45; DB 1; Length 289;
 Best Local Similarity 52.4%; Pred. No. 0.56;
 Matches 99; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 85 GCCCGGAGCCCAAGCAACGACGAAAGCAGGACGACGCGGGGTCTGGTCTTCCT 144
 |||||
 Db 268 GACCCGACCGGGAAAGACTCAAAAGTAAACCTTTTCCATCCAGGAACCTGGA 209
 |||||

QY 145 GGCTACAAGTACCTCGGACCCCTTCAACGGACTCGACAAAGGGGAGCCCGCTCAACGGCGG 204
 |||||
 Db 208 TCTCTGGGTATCACCGCGCAGCAGCTACAGCCCCAGGGGTGGAGCCCGCTCAGCCCTGTG 149
 |||||
 QY 205 GAGCAGCGGGCCCTCGAGCACGACAAAGGCTACGACACGACGCTCAAAAGCGGTGACAAT 264
 |||||
 Db 148 AGCTCACCAGTCTGACCACGACAAAGGGGCTCCCAAGCAGCTCGAAGAGCTCGACAA 89
 |||||
 QY 265 CGGTACCTTG 273
 |||||
 Db 88 ACCCACCTTG 80
 |||||

RESULT 13
 AW206887/c
 LOCUS
 DEFINITION
 UI-H-B11-afa-a-02-0-UI.s1 NCI CGAP_Sub3 Homo sapiens linear EST 02-DEC-1999
 IMAGE:2722826 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 322)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Oligo-dT track not found. Not I site shown in beginning of sequence is likely internal to the message. cDNA library preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 Location/Qualifiers
 1..322
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2722826"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Sub3"
 /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NCI CGAP Sub3 library is a subtracted library derived from the NCI CGAP Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI CGAP libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Br28, NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kid5, NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid11, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8, NCI CGAP CLL1, NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lu5, NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6, NCI CGAP Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
 NCI CGAP Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI CGAP Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI CGAP GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983,

1475592-1476743); NCI CGAP Pr22 pool 1 LLM 2457-2459,
2758-2759, 3062-3068 (IMAGE Clones 985608-986759,
1101192-1101959, 1217928-1220615); NCI CGAP Col0 pool 1
LLM 2644-2653, 2871-2872 (IMAGE Clones 1057416-1061255,
114584-1145351). Subtraction was performed as previously
described (Bonaldo, Lennon & Soares (1996): Normalization
and Subtraction: Two Approaches To Facilitate Gene
Discovery. Genome Research 6, 791-806.
TAG_TISSUE=lung
TAG_LIB=NCI CGAP_Lu5
TAG_SEQ=CAAC*

ORIGIN

Query Match 2.0%; Score 45; DB 1; Length 322;
Best Local Similarity 52.4%; Pred. No. 0.58; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 85 GCCCGAAGCCCAAGCCAAACAGCAAGAGGAGCGCGGGGGTGTGTTCTTCT 144
Db 267 GACCCGACCGGGAAGAGCTCAAGTAAACCTTTTCCATCCAGGAATCGAATCTCGT 208
QY 145 GGTACAGTACTCTGGACCTTCAAGGACTCGACAGGGGAGCCCTCAACGGCGG 204
Db 207 TCTTGGGTATACACGGCAGAGCTACAGCCCGAAGGGGTGGAGCCCTCAGCCCTGTG 148
QY 205 GACGAGCGGCGCTCGAGCAGCAGCAAGGCTTACGACCAAGCTCAAGCGGTGACAAAT 264
Db 147 AGCTACCCAGTCTGACCCAGCAAGAGGGGCTCCCAAGCACTGGAAGAGCTCGACAAG 88
QY 265 CGGTACCTG 273
Db 87 AGCCACCTG 79

RESULT 14

BG170384
LOCUS 602322706F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4426104 5',
DEFINITION mRNA sequence.
ACCESSION BG170384.1 GI:12677087
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 356)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10172 row: k column: 01
High quality sequence stop: 353.
Location/Qualifiers
1..356
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4426104"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 89"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dr primed."

FEATURES

source
1..451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2073412"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI CGAP CLL1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTCGGAGCCGCGCATGCTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into

Average insert size 1.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 2.0%; Score 45; DB 2; Length 356;
Best Local Similarity 52.4%; Pred. No. 0.6; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 85 GCCCGAAGCCCAAGCCAAACAGCAAGAGGAGCGCGGGGGTGTGTTCTTCT 144
Db 101 GACCCGACCGGGAAGAGCTCAAGTAAACCTTTTCCATCCAGGAATCGAATCTCGT 160
QY 145 GGTACAGTACTCTGGACCTTCAAGGACTCGACAGGGGAGCCCTCAACGGCGG 204
Db 161 TCTTGGGTATACACGGCAGAGCTACAGCCCGAAGGGGTGGAGCCCTCAGCCCTGTG 220
QY 205 GACGAGCGGCGCTCGAGCAGCAGCAAGGCTTACGACCAAGCTCAAGCGGTGACAAAT 264
Db 221 AGCTACCCAGTCTGACCCAGCAAGAGGGGCTCCCAAGCACTGGAAGAGCTCGACAAG 280
QY 265 CGGTACCTG 273
Db 281 AGCCACCTG 289

RESULT 15

A1439550/c
LOCUS A1439550.1 GI:4305107
DEFINITION Homo sapiens (human)
ACCESSION A1439550
VERSION EST.
KEYWORDS Homo sapiens
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 451)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 663 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 438.
Location/Qualifiers
1..451
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2073412"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone_lib="NCI CGAP CLL1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTCGGAGCCGCGCATGCTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into

FEATURES

source

THIS PAGE BLANK (USPTO)

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	2209	100.0	2211	3	US-09-807-802A-12	Sequence 12, Appl	
2	2209	100.0	4718	3	US-09-807-802A-1	Sequence 1, Appl	
3	2202.6	99.7	7447	3	US-10-216-870-11	Sequence 11, Appl	
4	2101.8	95.1	4683	3	US-09-807-802A-19	Sequence 19, Appl	
5	1798	81.4	1800	3	US-09-807-802A-14	Sequence 14, Appl	
6	1603	72.6	1605	3	US-09-807-802A-16	Sequence 16, Appl	
7	1458.6	66.0	4072	3	US-09-770-315-4	Sequence 4, Appl	
8	1458.6	66.0	4679	3	US-10-038-972A-12	Sequence 12, Appl	
9	1458.6	66.0	7557	3	US-09-770-315-3	Sequence 3, Appl	
10	1458.6	66.0	8698	3	US-09-770-315-2	Sequence 2, Appl	
11	1453.6	65.8	8179	3	US-09-438-268-5	Sequence 5, Appl	
12	1436.4	65.0	4680	2	US-08-254-358-1	Sequence 1, Appl	
13	1436.4	65.0	4680	2	US-08-475-391-1	Sequence 1, Appl	
14	1436.4	65.0	4680	2	US-08-709-609-1	Sequence 1, Appl	
15	1436.4	65.0	4680	6	PCT-US95-07178-1	Sequence 1, Appl	
16	1433.4	64.9	4681	3	US-09-807-802A-18	Sequence 18, Appl	
17	1408.2	63.7	4675	3	US-09-782-378A-1	Sequence 1, Appl	
18	1408.2	63.7	4675	3	US-09-782-378A-2	Sequence 2, Appl	
19	1241.8	56.2	8151	3	US-09-438-268-2	Sequence 2, Appl	
20	940	42.5	7744	3	US-10-216-870-14	Sequence 14, Appl	
21	933	42.2	7214	3	US-09-438-268-1	Sequence 1, Appl	
22	932	42.2	2208	3	US-09-532-594B-5	Sequence 5, Appl	
23	930.4	42.1	4767	3	US-09-532-594B-1	Sequence 1, Appl	
24	730.8	33.1	2307	3	US-09-533-427-7	Sequence 7, Appl	

QY 241 CAGAGCTCAAGCGGTTGACAAATCCGTACTCTGCGGTATAACCAAGCGCGAGTCTT 300
DB 241 CAGAGCTCAAGCGGTTGACAAATCCGTACTCTGCGGTATAACCAAGCGCGAGTCTT 300
QY 301 CAGGAGCGTCTCAAGAGAGATACGTCCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
DB 301 CAGGAGCGTCTCAAGAGAGATACGTCCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
QY 361 GCCAAGAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT 420
DB 361 GCCAAGAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT 420
QY 421 GGAAGAAACGTCGGTAGAGCAGTCGCCACAAAGACAGACTCTCTCTCGGGCATCGGC 480
DB 421 GGAAGAAACGTCGGTAGAGCAGTCGCCACAAAGACAGACTCTCTCTCGGGCATCGGC 480
QY 481 AAGACAGGCGCAGCAGCCCGCTTAAAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG 540
DB 481 AAGACAGGCGCAGCAGCCCGCTTAAAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG 540
QY 541 TCAGTCCCGATTCACAACTCTCGGAGAACCTTCAGCAACCCCGCTCTGTGGGACT 600
DB 541 TCAGTCCCGATTCACAACTCTCGGAGAACCTTCAGCAACCCCGCTCTGTGGGACT 600
QY 601 ACTACAAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGACGGA 660
DB 601 ACTACAAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGACGGA 660
QY 661 GTGGGTAAATCGCTCAGGAAATGGCAATTCGCAATTCACATGGCTGGCGCAGAGTCAATC 720
DB 661 GTGGGTAAATCGCTCAGGAAATGGCAATTCGCAATTCACATGGCTGGCGCAGAGTCAATC 720
QY 721 ACCACGACGCCGACCTTGGGCTTGGCCCACTTCAATTAACCACTCTTCAAGCAAAATC 780
DB 721 ACCACGACGCCGACCTTGGGCTTGGCCCACTTCAATTAACCACTCTTCAAGCAAAATC 780
QY 781 TCAGTGTCTCAACGGGGGCGCAGCAACGACCACTACTCTGGCTACAGCACCCCGCTGG 840
DB 781 TCAGTGTCTCAACGGGGGCGCAGCAACGACCACTACTCTGGCTACAGCACCCCGCTGG 840
QY 841 GGGTATTTTGAATTTCAACAGATTTCCACTCGCACTTTTACACGACTGGCAGGAGCTC 900
DB 841 GGGTATTTTGAATTTCAACAGATTTCCACTCGCACTTTTACACGACTGGCAGGAGCTC 900
QY 901 ATCAACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTTCTTCAACATCCAA 960
DB 901 ATCAACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTTCTTCAACATCCAA 960
QY 961 GTCAGGAGGTACGACGAATGATGGCGTCACAACTCGCTTAATTAACCTTACCAGCAG 1020
DB 961 GTCAGGAGGTACGACGAATGATGGCGTCACAACTCGCTTAATTAACCTTACCAGCAG 1020
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTCCTCGCTCTCGGCTCTGCGCACCCAG 1080
DB 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTCCTCGCTCTGCGCACCCAG 1080
QY 1081 GGTGCTCTCCCTCGCTTCCGGCGGACGTTTCATGATTCGGCAATACCGCTACCTGACG 1140
DB 1081 GGTGCTCTCCCTCGCTTCCGGCGGACGTTTCATGATTCGGCAATACCGCTACCTGACG 1140
QY 1141 CTCACAAATGGGAGCCAGCGGTGGGACGTTTCATCTTTTACTGCTGGAAATTTTCCCT 1200
DB 1141 CTCACAAATGGGAGCCAGCGGTGGGACGTTTCATCTTTTACTGCTGGAAATTTTCCCT 1200
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTTACACCTTTGAGGAGTGCCT 1260
DB 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTTACACCTTTGAGGAGTGCCT 1260
QY 1261 TTCCACAGCAGCTACCGCAGCAGCCAGAGCCTGGACCGGCTGTGATGATCTCTCATCGAC 1320
DB 1261 TTCCACAGCAGCTACCGCAGCAGCCAGAGCCTGGACCGGCTGTGATGATCTCTCATCGAC 1320

QY 1321 CAATACCTGTATTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGAC 1380
DB 1321 CAATACCTGTATTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGAC 1380
QY 1381 TTGCTGTTTACCGGTGCTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACT 1440
DB 1381 TTGCTGTTTACCGGTGCTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACT 1440
QY 1441 GGACCTGTTTATCGGACAGCGGCTTTCTAAAACAAAACAGACAAACACAGCAAT 1500
DB 1441 GGACCTGTTTATCGGACAGCGGCTTTCTAAAACAAAACAGACAAACACAGCAAT 1500
QY 1501 TTTTACTCGACTGGTCTTCAAAATATAAATTAACCTCAATGGCGGTGAATCATCATCAACCT 1560
DB 1501 TTTTACTCGACTGGTCTTCAAAATATAAATTAACCTCAATGGCGGTGAATCATCATCAACCT 1560
QY 1561 GGCACCTGTCTGCGCTCACACAAAGACGACGAAAGATTTTGGGACCGGTG 1620
DB 1561 GGCACCTGTCTGCGCTCACACAAAGACGACGAAAGATTTTGGGACCGGTG 1620
QY 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGCAATGTCTATGAT 1680
DB 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGCAATGTCTATGAT 1680
QY 1681 ACAGACGAGAGGAAATTAAGCCACTAACTGTGGCCACCGAAGATTTTGGGACCGGTG 1740
DB 1681 ACAGACGAGAGGAAATTAAGCCACTAACTGTGGCCACCGAAGATTTTGGGACCGGTG 1740
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACGACCTGCGACCGGAGATGTGCTATGGGA 1800
DB 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACGACCTGCGACCGGAGATGTGCTATGGGA 1800
QY 1801 GCATTAACCTGGCATGGTGTGCAAGATAGAGACGTGTACCTGACGGGTCCCAATTTGGGCG 1860
DB 1801 GCATTAACCTGGCATGGTGTGCAAGATAGAGACGTGTACCTGACGGGTCCCAATTTGGGCG 1860
QY 1861 AAAATTCCTCACACAGATGGACACTTTTCCCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920
DB 1861 AAAATTCCTCACACAGATGGACACTTTTCCCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920
QY 1921 AAGAACCCGCTCTCAGATCTCTCAATAAAACACGCTGTTCCTGCGAAATCTCCGCGG 1980
DB 1921 AAGAACCCGCTCTCAGATCTCTCAATAAAACACGCTGTTCCTGCGAAATCTCCGCGG 1980
QY 1981 GAGTTTTCAGCTAAGAGTTTGTCTTCAATCATACCAATATCTCCAGGACCAAGTGA 2040
DB 1981 GAGTTTTCAGCTAAGAGTTTGTCTTCAATCATCACCAATATCTCCAGGACCAAGTGA 2040
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAG 2100
DB 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAG 2100
QY 2101 TACACATCCAAATTTATGCAAAATCTGCCAACGTTTGAATTTTACTGTGGACAAATGGACT 2160
DB 2101 TACACATCCAAATTTATGCAAAATCTGCCAACGTTTGAATTTTACTGTGGACAAATGGACT 2160
QY 2161 TATACCTGAGCCTCGGCCCATTTGGACCCGCTTACCTTACCCGTCCTCTGT 2209
DB 2161 TATACCTGAGCCTCGGCCCATTTGGACCCGCTTACCTTACCCGTCCTCTGT 2209

RESULT 2

US-09-807-802A-1
; Sequence 1, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21


```

Db      6001  |||||TTGCTGTTTATGCGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACCTGGCTACCT 6060
Qy      1441  |||||GGACCTGTTATTCGGCAGCAGCGCTTCTTAAACCAAAACAGACAAACAAACGCAAT 1500
Db      6061  |||||GGACCTGTTATTCGGCAGCAGCGCTTCTTAAACCAAAACAGACAAACAAACGCAAT 6120
Qy      1501  |||||TTTACCTGGACTGGTCTTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCT 1560
Db      6121  |||||TTTACCTGGACTGGTCTTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCT 6180
Qy      1561  |||||GGCAGTGTATGGCTCTCACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTC 1620
Db      6181  |||||GGCAGTGTATGGCTCTCACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTC 6240
Qy      1621  |||||ATGATTTTTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATT 1680
Db      6241  |||||ATGATTTTTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATT 6300
Qy      1681  |||||ACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
Db      6301  |||||ACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGTG 6360
Qy      1741  |||||GCAGTCAATTCAGAGCAGCAGCAGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGA 1800
Db      6361  |||||GCAGTCAATTCAGAGCAGCAGCAGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGA 6420
Qy      1801  |||||GCATTAACCTGGCATGTGGCAAGATAGAGAGCTACTGACAGGTCCTCAATTTGGGCC 1860
Db      6421  |||||GCATTAACCTGGCATGTGGCAAGATAGAGAGCTACTGACAGGTCCTCAATTTGGGCC 6480
Qy      1861  |||||AAAAATTCCTCACACAGATGGACACTTTTCAACCCCTCTCTTATATGGCGGCTTTGGACTC 1920
Db      6481  |||||AAAAATTCCTCACACAGATGGACACTTTTCAACCCCTCTCTTATATGGCGGCTTTGGACTC 6540
Qy      1921  |||||AAGAACCCGCTCTCAGATCTCATCAATAAACAACGCGCTGTTCTCGGAATCCTCCGGCG 1980
Db      6541  |||||AAGAACCCGCTCTCAGATCTCATCAATAAACAACGCGCTGTTCTCGGAATCCTCCGGCG 6600
Qy      1981  |||||GAGTTTTTCACTACAAGTTTGTCTTCAATCACCCTAATCTCCACAGACAAAGTGAGT 2040
Db      6601  |||||GAGTTTTTCACTACAAGTTTGTCTTCAATCACCCTAATCTCCACAGACAAAGTGAGT 6660
Qy      2041  |||||GTGGAATTAATGGAGCTGCAAGAAAGAAACAGACGCGCTGGAAATCCCGAAGTGCAG 2100
Db      6661  |||||GTGGAATTAATGGAGCTGCAAGAAAGAAACAGACGCGCTGGAAATCCCGAAGTGCAG 6720
Qy      2101  |||||TACACATCCAATTAATGCAAAATCTGCCAACGTTTGTATTTTACTGTGGACAAATGAGATT 2160
Db      6721  |||||TACACATCCAATTAATGCAAAATCTGCCAACGTTTGTATTTTACTGTGGACAAATGAGATT 6780
Qy      2161  |||||TATACGTAGCTCGCCCATTTGGCAGCCGTTACTTACCCGTCCTCTGT 2209
Db      6781  |||||TATACGTAGCTCGCCCATTTGGCAGCCGTTACTTACCCGTCCTCTGT 6829
```

```

RESULT 4
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: RAV-6
; US-09-807-802A-19

Query Match      95.1%; Score 2101.8; DB 3; Length 4683;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2142; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy      1  ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAAACTCTCTGAGGCGCATTCG 60
Db      2208  ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAAACTCTCTGAGGCGCATTCG 2267
Qy      61  GAGTGTGGGACTTCAAACTCGAGCCCGAAAGCCCAAGGCCAAACCCAGCAAAAGCAGGAC 120
Db      2268  GAGTGTGGGACTTCAAACTCGAGCCCGAAAGCCCAAGGCCAAACCCAGCAAAAGCAGGAC 2327
Qy      121  GACGCGCGGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC 180
Db      2328  GACGCGCGGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC 2387
Qy      181  AAGGGGGAGCCCGTCAACGCGGGGAGCCGAGCGGCTCGAGCAGCAGCAGGCTACGAC 240
Db      2388  AAGGGGGAGCCCGTCAACGCGGGGAGCCGAGCGGCTCGAGCAGCAGCAGGCTACGAC 2447
Qy      241  CAGCAGCTCAAAAGCGGGTGACAATCCGTACTCGGTATAAACCCAGCCGACGCGGAGTTT 300
Db      2448  CAGCAGCTCAAAAGCGGGTGACAATCCGTACTCGGTATAAACCCAGCCGACGCGGAGTTT 2507
Qy      301  CAGGAGCGTCTCAAGAAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360
Db      2508  CAGGAGCGTCTCAAGAAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 2567
Qy      361  GCCAAGAACGGGTTCTCGAACTCTCGGTCTGGTTGAGGAGGCGCTTAAGCGGCTCCT 420
Db      2568  GCCAAGAACGGGTTCTCGAACTTTTGGTCTGGTTGAGGAGGCTTAAGCGGCTCCT 2627
Qy      421  GGAAAGAAACGTCGGTAGAGCAGTCCGCAAGAGCCAGACTCTCTCGGCGCATCGGC 480
Db      2628  GGAAAGAAACGTCGGTAGAGCAGTCCGCAAGAGCCAGACTCTCTCGGCGCATCGGC 2687
Qy      481  AAGACAGGCGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAGACTGCGCAGTCAGAG 540
Db      2688  AAGACAGGCGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAGACTGCGCAGTCAGAG 2747
Qy      541  TCAGTCCCGCATCCAAACCTCTCGGAGAACCTTCAGAGAACCCCGCTGCTGGGACCT 600
Db      2748  TCAGTCCCGCATCCAAACCTCTCGGAGAACCTTCAGAGAACCCCGCTGCTGGGACCT 2807
Qy      601  ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAAATAACGAAGCGCGCAGCGA 660
Db      2808  ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAAATAACGAAGCGCGCAGCGA 2867
Qy      661  GTGGGTAAATGCTCTCAGGAAATTTGGCATTTCCATGCTGGGCGCAGAGATCATC 720
Db      2868  GTGGGTAAATGCTCTCAGGAAATTTGGCATTTCCATGCTGGGCGCAGAGATCATC 2927
Qy      721  ACCACGACGACCCGCACTTGGGCTTGGCCACCTACAATACCACTCTTACAGCAATC 780
Db      2928  ACCACGACGACCCGCACTTGGGCTTGGCCACCTACAATACCACTCTTACAGCAATC 2987
Qy      781  TCCAGTGTCTTCAACGGGGCGCAGCAACCACTACTTCTGGGTACAGCAGCCCTCGG 840
Db      2988  TCCAGTGTCTTCAACGGGGCGCAGCAACCACTACTTCTGGGTACAGCAGCCCTCGG 3047
Qy      841  GGGTATTTTGAATTTCAACAGATTTCCACTGCGCACTTTTTCACCACTGCTGGCAGCGACTC 900
Db      3048  GGGTATTTTGAATTTCAACAGATTTCCACTGCGCACTTTTTCACCACTGCTGGCAGCGACTC 3107
Qy      901  ATCAACAACAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
Db      |||||
```

```
Db 3108 ATCAACAACTTGGGATTTCGGCCCAAGAGACTCAACTTCAAGCTCTTCAACATCCAA 3167
Qy 961 GTCAAGGAGGTACAGCAAGTAATGATGGCTGCAACCATCGCTAAATAACCTTACAGACAG 1020
Db 3168 GTCAAGGAGGTACAGCAAGTAATGATGGCTGCAACCATCGCTAAATAACCTTACAGACAG 3227
Qy 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTTCGTAAGTCTCGGCTCTCGGACACAG 1080
Db 3228 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTTCGTAAGTCTCGGCTCTCGGACACAG 3287
Qy 1081 GGCTGCTCCCTCCGTTCCGGCGGACAGTGTTCATGATTCGCAATACAGCTTACCTGACG 1140
Db 3288 GGCTGCTCCCTCCGTTCCGGCGGACAGTGTTCATGATTCGCAATACAGCTTACCTGACG 3347
Qy 1141 CTCAACAATGGGAGCCCAAGCCGTGGGACGTTTCACTTTTACTGCTTGGAAATATTTCCTT 1200
Db 3348 CTCAACAATGGGAGCCCAAGCCGTGGGACGTTTCACTTTTACTGCTTGGAAATATTTCCTT 3407
Qy 1201 TCTCAGATGCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCT 1260
Db 3408 TCGCAGATGCTGAGAACGGGCAATTAATTTTACCTTCAGCTACACCTTCGAGGACGTGCT 3467
Qy 1261 TTCCACAGAGCTACGCGACAGCAGAGCTGGACCGGCTGATGAATCTCTCATCGAC 1320
Db 3468 TTCCACAGAGCTACGCGACAGCAGAGCTGGACCGGCTGATGAATCTCTCATCGAC 3527
Qy 1321 CAATACCTGTATTACTGGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGAC 1380
Db 3528 CAGTACCTGTATTACTGGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGAC 3587
Qy 1381 TTGCTGTTTGTAGCGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACCT 1440
Db 3588 TTGCTGTTTGTAGCGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACCT 3647
Qy 1441 GGACCTGTATTTCGGCAGCAGCGCGTTTCTAAAAAAGAGAGCAAGTCTTTTCCATGAGCGGTGC 1500
Db 3648 GGACCTGTATTTCGGCAGCAGCGCGTTTCTAAAAAAGAGAGCAAGTCTTTTCCATGAGCGGTGC 3707
Qy 1501 TTTTACCTGAGCTGGTCTTCAAAATATAAATCCTCAATGGCGGTGAATTCATCATCAACCTT 1560
Db 3708 TTTTACCTGAGCTGGTCTTCAAAATATAAATCCTCAATGGCGGTGAATTCATCATCAACCTT 3767
Qy 1561 GGCACCTGTATGGCTTCACAAAGACGAGAGCAAGTCTTTTCCATGAGCGGTGC 1620
Db 3768 GGCACCTGTATGGCTTCACAAAGACGAGAGCAAGTCTTTTCCATGAGCGGTGC 3827
Qy 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAATGTCAATGAT 1680
Db 3828 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAATGTCAATGAT 3887
Qy 1681 ACAGACGAAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
Db 3888 ACAGACGAAGAGAAATCAAAAGCACTAAACCCGTGGCCACCGAAAGATTTGGGACCTGTG 3947
Qy 1741 CGAGTCAATTTCCAGAGCAGCAGACACCTTCGACCGGAGATGTGATGCTATGGGA 1800
Db 3948 CGAGTCAATTTCCAGAGCAGCAGACACCTTCGACCGGAGATGTGATGCTATGGGA 4007
Qy 1801 GCATTAACCTGGCATGTGTGGCAAGATAGAGAGCTGTACCTTCAGCGGTGCCATTTGGGCG 1860
Db 4008 GCCTTAACCTGGCATGTGTGGCAAGATAGAGAGCTGTACCTTCAGCGGTGCCATTTGGGCG 4067
Qy 1861 AAAATTCCTCACACAGATGGACACTTTTCACCCGTCTCTTATTTGGGCGGCTTTGGACTC 1920
Db 4068 AAAATTCCTCACACAGATGGACACTTTTCACCCGTCTCTTATTTGGGCGGCTTTGGACTT 4127
Qy 1921 AGAAGCCCGCTCTCAGATCTCTCATCAAAACACAGCCCTGTTCTCCGAAATCTCCCGCG 1980
Db 4128 AGAAGCCCGCTCTCAGATCTCTCATCAAAACACAGCCCTGTTCTCCGAAATCTCCCGCG 4187
Qy 1981 GAGTTTTTACAGTACAAAGTTTGTCTTCAATTCATCACCCTAATCTCCACAGCAAGTGAGT 2040
Db 4188 GAGTTTTTACAGTACAAAGTTTGTCTTCAATTCATCACCCTAATCTCCACAGCAAGTGAGC 4247
```

```
Qy 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGACAGCGCTGGAATCCCGAAGTGCAG 2100
Db 4248 GTGGAGATTTGAATGGGAGCTGCAGAAAGAAACAGCAACCGTGGAAATCCCGAAGTGCAG 4307
Qy 2101 TACACATCCAATTTATGCAAAATCTGCCAACTGTGATTTTACTGTGGACAAACAATGGACTT 2160
Db 4308 TATACATCTAATTTATGCAAAATCTGCCAACTGTGATTTTACTGTGGACAAACAATGGACTT 4367
Qy 2161 TATACTAGCTTCGCGCCCAATTTGGCACCCGTTTACCTTACCTTCCTCCCTCT 2209
Db 4368 TATACTAGCTTCGCGCCCAATTTGGCACCCGTTTACCTTACCTTCCTCCCTCT 4416
```

RESULT 5

```
US-09-807-802A-14
; Sequence 14, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1797)
; OTHER INFORMATION:
US-09-807-802A-14
```

```
Query Match 81.4%; Score 1798; DB 3; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 412 ACGGCTCTCTGGAAGAAACGTCGCTAGAGCAGTGCACCAAGAGCCAGACTCTCTCTCG 471
Db 1 ACGGCTCTCTGGAAGAAACGTCGCTAGAGCAGTGCACCAAGAGCCAGACTCTCTCTCG 60

Qy 472 GGCATCGGCAAGACAGGCGCAGCGCTAAAAGAGACTCAATTTTGGTCAGACTGGC 531
Db 61 GGCATCGGCAAGACAGGCGCAGCGCTAAAAGAGACTCAATTTTGGTCAGACTGGC 120

Qy 532 GACTCAGAGTCACTCCCGATCCCAACCTCTCGGAGACCTCCAGCAACCCCGCTGT 591
Db 121 GACTCAGAGTCACTCCCGATCCCAACCTCTCGGAGACCTCCAGCAACCCCGCTGT 180

Qy 592 GTGGGACCTTACTTCAATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGC 651
Db 181 GTGGGACCTTACTTCAATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGC 240

Qy 652 GCCGACGAGTGGTAAATGGCTCAGGAAATTTGGGATTTGGATTCACATGGCTGGCGAC 711
Db 241 GCCGACGAGTGGTAAATGGCTCAGGAAATTTGGGATTTGGATTCACATGGCTGGCGAC 300

Qy 712 AGAGTCATCACCACAGCAGCCGACCTTGGCGCTTGGCCACCTCAATAACCACTCTAC 771
Db 301 AGAGTCATCACCACAGCAGCCGACCTTGGCGCTTGGCCACCTCAATAACCACTCTAC 360

Qy 772 AAGCAAAATCTCAGTGTCTTCAACGGGGGCGCAGCAACCACTTCTCGGCTACAGC 831
Db 361 AAGCAAAATCTCAGTGTCTTCAACGGGGGCGCAGCAACCACTTCTCGGCTACAGC 420
```

QY 832 ACCCCCTGGGGTATTTTGAATTTCAACAGATTCCTCACTGCCACTTTTCCACCAAGTGAATG 891
DB 421 ACCCCCTGGGGTATTTTGAATTTCAACAGATTCCTCACTGCCACTTTTCCACCAAGTGAATG 480
QY 892 CAGCGACTCATCAACAACTTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTC 951
DB 481 CAGCGACTCATCAACAACTTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTC 540
QY 952 AACATCAAGTCAAGGAGTCAAGCAATGATGGGCTCACAACCATCGCTAATACCTT 1011
DB 541 AACATCAAGTCAAGGAGTCAAGCAATGATGGGCTCACAACCATCGCTAATACCTT 600
QY 1012 ACCAGCAGCTTCAAGTCTTCTGGACTCGGAGTACCAGTTCCTGAGTCTCTCGGCTCT 1071
DB 601 ACCAGCAGCTTCAAGTCTTCTGGACTCGGAGTACCAGTTCCTGAGTCTCTCGGCTCT 660
QY 1072 GGGCAGCAGGCTGCTCCTCGTTCGGGGGAGCTGTTCTATGATTCGGCAATACGGC 1131
DB 661 GGGCAGCAGGCTGCTCCTCGTTCGGGGGAGCTGTTCTATGATTCGGCAATACGGC 720
QY 1132 TACCTGACCTCAACAACTGCGAGCCAGCGCTGGGAGCTTCATCTTTTACTGCTGGAA 1191
DB 721 TACCTGACCTCAACAACTGCGAGCCAGCGCTGGGAGCTTCATCTTTTACTGCTGGAA 780
QY 1192 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTGAG 1251
DB 781 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTGAG 840
QY 1252 GAAGTGCCTTTTCCACAGCAGTACGCGCAGCAGAGCTTGACCGGCTGATGAATCT 1311
DB 841 GAAGTGCCTTTTCCACAGCAGTACGCGCAGCAGAGCTTGACCGGCTGATGAATCT 900
QY 1312 CTCATCGACCAATACCTGATTTACCTGACAGAACTCAAAATCAGTCCGGAAGTGCCAA 1371
DB 901 CTCATCGACCAATACCTGATTTACCTGACAGAACTCAAAATCAGTCCGGAAGTGCCAA 960
QY 1372 AACAAAGACTTGTGTTAGCCGTGGGTCTCAGCTGGGATGCTGTTGAGCCCAAAAC 1431
DB 961 AACAAAGACTTGTGTTAGCCGTGGGTCTCAGCTGGGATGCTGTTGAGCCCAAAAC 1020
QY 1432 TGCTACTTGGACCTTGTATCGGAGAGCGCGTTTCAAAACAAACAAACAAAC 1491
DB 1021 TGCTACTTGGACCTTGTATCGGAGAGCGCGTTTCAAAACAAACAAACAAAC 1080
QY 1492 AACAGCAATTTTACCTGGAGTGTCTTCAAAATATAAATCAATGAGGCTGATCCATC 1551
DB 1081 AACAGCAATTTTACCTGGAGTGTCTTCAAAATATAAATCAATGAGGCTGATCCATC 1140
QY 1552 ATCAACCTGGCACTGCTATGGCTTCACAAAGAGAGCGCAAGCAAGTTCCTTCCCATG 1611
DB 1141 ATCAACCTGGCACTGCTATGGCTTCACAAAGAGAGCGCAAGCAAGTTCCTTCCCATG 1200
QY 1612 AGCGGTGTCATGATTTTGGAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAT 1671
DB 1201 AGCGGTGTCATGATTTTGGAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAT 1260
QY 1672 GTCATGATTTACAGCAAGGAGAAATTAAGCCACTAAACCTGAGCCACCGAAGATTT 1731
DB 1261 GTCATGATTTACAGCAAGGAGAAATTAAGCCACTAAACCTGAGCCACCGAAGATTT 1320
QY 1732 GGGACGCTGGCAGTCAATTTCCAGAGCAGCAGCAGACCCCTCGACCGGAGATGTCAT 1791
DB 1321 GGGACGCTGGCAGTCAATTTCCAGAGCAGCAGCAGACCCCTCGACCGGAGATGTCAT 1380
QY 1792 GCTATGGGAGCAATTAACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAAGGTCCC 1851
DB 1381 GCTATGGGAGCAATTAACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAAGGTCCC 1440
QY 1852 ATTTGGGCCAAAATTTCTCTACAGATGGAACATTTTCAACCCGCTCTCTTATGGGGGGC 1911
DB 1441 ATTTGGGGCCAAAATTTCTCTACAGATGGAACATTTTCAACCCGCTCTCTTATGGGGGGC 1500

QY 1912 TTTGAGCTCAAGAACCCGCTCCTCAGATCTCTATCAAAAAACACGCTCTTCTCTCGGAAT 1971
DB 1501 TTTGAGCTCAAGAACCCGCTCCTCAGATCTCTATCAAAAAACACGCTCTTCTCTCGGAAT 1560
QY 1972 CCTCCGGCGGAGTTTTCAGCTACAAAGTTTCTTATTCATCAACCAATCTCCACAGGA 2031
DB 1561 CCTCCGGCGGAGTTTTCAGCTACAAAGTTTCTTATTCATCAACCAATCTCCACAGGA 1620
QY 2032 CAAAGTGAGTGTGGAAATTTGAATGGGAGCTGCGAGAAAGAAAAACAGCAAGCGCTGGAAATCCC 2091
DB 1621 CAAAGTGAGTGTGGAAATTTGAATGGGAGCTGCGAGAAAGAAAAACAGCAAGCGCTGGAAATCCC 1680
QY 2092 GAAGTGCAGTACACATCCAAATTTATGCAAAATCTGCAAGTTTATTTTACTGTGGACAAC 2151
DB 1681 GAAGTGCAGTACACATCCAAATTTATGCAAAATCTGCAAGTTTATTTTACTGTGGACAAC 1740
QY 2152 AATGAGCTTTTACTGAGCTCGCCCATTTGGCACCCTTACTTACCCGCTCCCTGT 2209
DB 1741 AATGAGCTTTTACTGAGCTCGCCCATTTGGCACCCTTACTTACCCGCTCCCTGT 1798

RESULT 6
US-09-807-802A-16
; Sequence 16, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 1605
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1602)
; OTHER INFORMATION:
US-09-807-802A-16

Query Match 72.6%; Score 1603; DB 3; Length 1605;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 607 ATGGCTTTCAGGCGGTGGCGCAACAATGGCAGACAATAAGAGGCGCGCAGAGTGGGT 666
DB 1 ATGGCTTTCAGGCGGTGGCGCAACAATGGCAGACAATAAGAGGCGCGCAGAGTGGGT 60
QY 667 AATGCTTCAGGAAATTTGGCAATTCGATTCACATGGCTGGCGCAGAGTTCATCACCACC 726
DB 61 AATGCTTCAGGAAATTTGGCAATTCGATTCACATGGCTGGCGCAGAGTTCATCACCACC 120
QY 727 AGCACCAGCAGCTGGGCTTGGCCACCTTAACTAACTAACTTAAAGCAAAATCTCCAGT 786
DB 121 AGCACCAGCAGCTGGGCTTGGCCACCTTAACTAACTAACTTAAAGCAAAATCTCCAGT 180
QY 787 GCTTCAACGGGGCCAGCAACCACTACTTCTCGGTACAGACCCCTTGGGGGTAT 846
DB 181 GCTTCAACGGGGCCAGCAACCACTACTTCTCGGTACAGACCCCTTGGGGGTAT 240
QY 847 TTTGATTTCAACAGATTCCTGCTGCTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
DB 241 TTTGATTTCAACAGATTCCTGCTGCTTTCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

QY 907 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAAGTCAAG 966
DB 301 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAAGTCAAG 360
QY 967 GAGGTACAGACGAATGATGGCGTCAACACCATCGCTAATAACCTTTACAGACACGGTTCAA 1026
DB 361 GAGGTACAGACGAATGATGGCGTCAACACCATCGCTAATAACCTTTACAGACACGGTTCAA 420
QY 1027 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCTCGGTCTCGGCTCGGACACGGGCTGC 1086
DB 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCTCGGTCTCGGCTCGGACACGGGCTGC 480
QY 1087 CTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGGTCTACCTGACGCTCAAC 1146
DB 481 CTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGGTCTACCTGACGCTCAAC 540
QY 1147 AATGGCAGCAAGCCGTCGGAGCGTTTCATCTCTTTTACTGCTCGGAATATTTCCCTTCTCAG 1206
DB 541 AATGGCAGCAAGCCGTCGGAGCGTTTCATCTCTTTTACTGCTCGGAATATTTCCCTTCTCAG 600
QY 1207 ATGCTGAGAACCGGGAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCTTTCAC 1266
DB 601 ATGCTGAGAACCGGGAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCTTTCAC 660
QY 1267 AGCAGTACGCGACAGCAGACGCTTGACCGGCTGATGAATCCTCTCATCGACCAATAC 1326
DB 661 AGCAGTACGCGACAGCAGACGCTTGACCGGCTGATGAATCCTCTCATCGACCAATAC 720
QY 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAAAGGACTTGTG 1386
DB 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAAAGGACTTGTG 780
QY 1387 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCGGACCC 1446
DB 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCGGACCC 840
QY 1447 TGTATTCGCGACGCGCGTCTTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 1506
DB 841 TGTATTCGCGACGCGCGTCTTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 900
QY 1507 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 1566
DB 901 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960
QY 1567 GCTATGGCTCACAACAAAGACGACGAAGTCTTTCCATGAGCGGTGTGATGATT 1626
DB 961 GCTATGGCTCACAACAAAGACGACGAAGTCTTTCCATGAGCGGTGTGATGATT 1020
QY 1627 TTTGGAABAGAGAGCGCGGAGCTTCAACACTGCAATTCGACAAATGTCTATGATTACAGAC 1686
DB 1021 TTTGGAABAGAGAGCGCGGAGCTTCAACACTGCAATTCGACAAATGTCTATGATTACAGAC 1080
QY 1687 GAAGAGGAATTAAGGCCACTAAACCTGTGGCCACCGAAAGATTTTGGGACCGTGGCAGTTC 1746
DB 1081 GAAGAGGAATTAAGGCCACTAAACCTGTGGCCACCGAAAGATTTTGGGACCGTGGCAGTTC 1140
QY 1747 AATTTCCAGACGACGACGACAGACCTCTGCAACCGGAGATGTGATGCTATGAGGACATTA 1806
DB 1141 AATTTCCAGACGACGACGACAGACCTCTGCAACCGGAGATGTGATGCTATGAGGACATTA 1200
QY 1807 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAATTT 1866
DB 1201 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAATTT 1260
QY 1867 CTTTCAACAGATGGACACTTTTCAACCGTCTCTCTTTATGGCGGCTTTTGGACTCAAGAAT 1926
DB 1261 CTTTCAACAGATGGACACTTTTCAACCGTCTCTCTTTATGGCGGCTTTTGGACTCAAGAAT 1320
QY 1927 CGGCTCTCTCAGATCTCTCAATAAACAACGCTGTCTCTGCGAATCTCTCGGCGGAGTTT 1986
DB 1321 CGGCTCTCTCAGATCTCTCAATAAACAACGCTGTCTCTGCGAATCTCTCGGCGGAGTTT 1380
QY 1987 TCAGCTACAAAGTTTGTCTTCAATCATCAACCCATCTCTCCACAGACAAAGTGTGTGGAA 2046

DB 1381 TCAGCTACAAAGTTTGTCTTCAATCATCAACCAATCTCCACAGACAAAGTGTGTGGAA 1440
QY 2047 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 2106
DB 1441 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACACA 1500
QY 2107 TCCAAATTATGCAAAATCTGCCAAGCTTGATTTTACTGTGGACAAATGGACTTTTACT 2166
DB 1501 TCCAAATTATGCAAAATCTGCCAAGCTTGATTTTACTGTGGACAAATGGACTTTTACT 1560
QY 2167 GAGCTCTGCCCATTTGGCACCCGTTACCTTACCCTGCCCTGT 2209
DB 1561 GAGCTCTGCCCATTTGGCACCCGTTACCTTACCCTGCCCTGT 1603

RESULT 7

US-09-770-315-4

; Sequence 4, Application US/09770315

; Patent No. 6429001

; GENERAL INFORMATION:

; APPLICANT: Chiron Corporation

; TITLE OF INVENTION: Recombinant AAV Packaging Systems

; FILE REFERENCE: 20263-501

; CURRENT APPLICATION NUMBER: US/09/770,315

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: US 60/178,536

; PRIOR FILING DATE: 2000-01-26

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 4072

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: recombinant DNA

US-09-770-315-4

Query Match 66.0%; Score 1458.6; DB 3; Length 4072;

Best Local Similarity 79.1%; Pred. No. 0;

Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;

QY 1 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60
DB 1484 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTGAGGAATAAGA 1543
QY 61 GAGTGTGGGACTTGAACCTTGAGCCCGGAGCCCAAGCCCAACAGCAAAAGCAGAC 120
DB 1544 CAGTGTGGAAGCTCAAACTTGGCCCAACACCAAGCCGAGCGGCTAAGGAC 1603
QY 121 GACGCGCGGGTCTGGTCTTCTGGCTTCAAGTACCTCGGACCTTCAACGGACTCGAC 180
DB 1604 GACAGCAGGGGTCTTGTCTTCTGGGTACAAGTACCTCGGACCTTCAACGGACTCGAC 1663
QY 181 AAGGGGAGCCGCTCAACGCGGACCGAGCCGCGCTCGAGCAGCAAGCCCTACGAC 240
DB 1664 AAGGGAGAGCCGCTCAACGAGCAGCCGCGCCCTCGAGCAGCAAGCCCTACGAC 1723
QY 241 CAGCAGCTCAAAAGCGGGTGACAACTCCGTACCTGCGGTATAACCAACGCGAGCCGAGTTT 300
DB 1724 CGGCAGCTCGACGCGGAGACAAACCGTACCTCAAGTACAAACCAACGCGGAGTTT 1783
QY 301 CAGGAGCGTCTGCAAGAGATACGCTTTTGGGGGCAACCTCGGCGGAGCAGTCTTCCAG 360
DB 1784 CAGGAGCGCTTAAAGAGATACGCTTTTGGGGGCAACCTCGGCGGAGCAGTCTTCCAG 1843
QY 361 GCCAAGAGCGGTTCTCGAACCCTCTCGTCTGGTTGAGGAGGCGCTAAAGACGGCTCT 420
DB 1844 GCGAAAAGAGAGGTTCTTGAACCTCTGGGCTTGGTTGAGGAACCTGTTAAGACGGCTCG 1903
QY 421 GGAAAGAAACGTCGCGTAGAGCAGTCGCAACAGAGCCAGACTCTCTCTCGGCGATCGAC 480
DB 1904 GGAAAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCTCGGAAACCGGA 1963

481 AAGACAGCCAGCAGCCGCTAAAAAGAGACTCAATTTTGGTGCAGACTGGCGACTCAGAG 540
1964 AAGCGCGCCAGCAGCCTGCAAGAAAAGATTGAATTTTGGTGCAGACTGGAGACCGAGAC 2023
541 TCAGTCCCGCATCCACAACCTCTCGGAGAACTCCAGCAACCCCGCTGCTGTGGACT 600
2024 TCAGTACCTGACCCCGCAGCTCTCGGACAGCCAGCAGCCCTCTGCTGTGGAACT 2083
601 ACTACAATGGCTTCAGCGGTGGCCGACCAATGGGAGACAATAAGAGGCGCCGACGGA 660
2084 AATACGATGGCTACAGCGAGTGGCGACCAATGGGAGACAATAAGAGGCGCCGACGGA 2143
661 GTGGGTAAATCCCTCAGGAAATTTGGCAATTCGCAATTCACATGGCTGGGCGACAGATCATC 720
2144 GTGGGTAAATTCCTCGGAAATTTGGCAATTCGCAATTCACATGGTGGGCGACAGATCATC 2203
721 ACCACAGACCCGCACTCGGCGCTTGGCCCACTCAATTAACCACTCTCAAGGCAATC 780
2204 ACCACAGACCCGCACTCGGCGCTTGGCCCACTCAATTAACCACTCTCAAGGCAATC 2263
781 TCCAGTGTCTCAACGGGGCCAGCAGACGACCACTACTTCCGCTACAGCAGCCCTGG 840
2264 TCCAGGCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCAGCCCTTGG 2320
841 GGGTATTTGATTTCAACAGATTTCCACTGCCACTTTTCAACCAGTGACTGGCAGGACTC 900
2321 GGGTATTTGACTTTCAACAGATTTCCACTGCCACTTTTCAACCAGTGACTGGCAGGACTC 2380
901 ATCAACAACAAATTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTTTCAACATCAA 960
2381 ATCAACAACAAATTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTTTCAACATCAA 2440
961 GTCAAGGAGTCAACGAGTATGATGGCGTCACAACTCGCTTAATACCTTACCGACGAG 1020
2441 GTCAAGGAGTCAACGAGTATGATGGCGTCACAACTCGCTTAATACCTTACCGACGAG 2500
1021 GTTCAAGTCTTCTCGGACTCGAGTACCACTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 1080
2501 GTTCAAGTCTTCTCGGACTCGAGTACCACTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 2560
1081 GGCTGCTCCCTCCGTTCCGCGGAGAGTGTTCATGATTTCCGAAATACCGCTACCTGAGC 1140
2561 GGATGCTCTCCGCGTTCCGAGCAGAGCTTTTCAATGGTGCCACAGTATGGATACCTCACC 2620
1141 CTCACAATGGGAGCCAAAGCGTGGAGCGTTCACTCTTTTACTCGCTGGAAATATTTCCCT 1200
2621 CTGAACAACGGGAGTCAGGAGTAGGACGCTCTTCAATTTTACTGCTGGAGTACTTTCT 2680
1201 TCTCAGATCTGAGAACGGGCAACAACTTTTACCTTCAGCTTACACCTTTGAGGAAGTGCCT 1260
2681 TCTCAGATCTGAGTAACGGGAAACAATTTTACCTTCAGCTTACACCTTTGAGGAGTGCCT 2740
1261 TTCCACAGAGCTACCGGCACAGCCAGAGCTGGACCGCTGATGATGATCTCTCATCGAC 1320
2741 TTCCACAGAGCTACGCTCAGACCGAGAGTCTGGACCGCTCTCATGATCTCTCATCGAC 2800
1321 CAATACCTGTATTTACCTGAACAGAACTCAAAATCAGTCCGGAGTCCCAACAAACAGGAC 1380
2801 CAGTACCTGTATTTACTTGAGCAGAAACAAACACTTCAAGTGGAAACCAACGAGTCAAGG 2860
1381 TTGCTGTTTACGCTGGTCTCCAGCTGGCATGCTGTTTACGCGCAAAAACCTGGCTACCT 1440
2861 CTTTCACTTTTCTCAGCGCGGAGGAGTGACATTTCCGGGACGAGTCTAGGAACTGGCTTCC 2920
1441 GGACCTGTTTATCGGAGCAGCGGCTTTTCAAAACAAAACAGACAAACAAACAGCAAT 1500
2921 GGACCTGTTTATCGGAGCAGCGGAGTATCAAGACATCTGCGGATTAACAAACAGAGTAA 2980
1501 TTTTACTGGAGCTGGTGTCAAAATATAACCTCAATGGGGGTGAATTCATCATCAACCT 1560
2981 TACTCTGGAGCTGGAGCTACCAAGTACCACTCAATGGCAGAGACTCTCTGTGGTGAATCCG 3040

1561 GGCACGTGTATGGCCTTCAACAAGACGACGAAGCAAGATTTTCCCATGAGCGGTGTC 1620
3041 GGCCTGGCCATGGCAAGCCACAAGAGCAGATGAAGAAAAATTTTCTCTCAGAGCGGGTT 3100
1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAACACTGCAATTTGGACAATGTCATGATT 1680
3101 CTATCTTTTGGAGCAAGGCTCAGAGAAAACAATTTGGACATTTGAAAAAGTCAATGATT 3160
1681 ACAGACGAAGAGGAAATTTAAAGCCACTAACCTGTGGCCACCGAAAGATTTTGGGACCGTG 1740
3161 ACAGACGAAGAGGAAATCAGGACACCAATCCCGTGGCTACGGAGCAGTATGGTCTGTA 3220
1741 GCAATCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCACTCATGGA 1800
3221 TCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACAAGGC 3280
1801 GCATTTACCTGGCATGCTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCAATTTGGGCG 1860
3281 GTTCTTCCAGGATGGTCTGGCAGCAGAGATGTGTACCTTTCAGGGGCCCATCTGGGCA 3340
1861 AAAATTTCTCTCACACAGATGGACATTTTCAACCGCTCTCTCTTTATGGGCGCTTTGACTC 1920
3341 AGATTTCCACACAGCAGCGACATTTTCAACCCCTCTCCCTCATGGTGGATTCGACTT 3400
1921 AAGAACCCGCTCTCAGATCTCATCAAAAACAGCCCTGTTCTCGGAACTCTCCGGC 1980
3401 AAACACCCCTCTCCACAGATTTCTCATCAAGAACACCCCGGTACCTGCGAATCTTTCGACC 3460
1981 GAGTTTTTCAGCTACAAAGTTTGTCTTCAATTCACCACTACTCCACAGGACAAAGTGA 2040
3461 ACCTTCAGTGGCGGAAAGTTTGTCTTCTTCAATCACAGTACTCCACGGGACAGTCA 3520
2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAAAACAGCAAGCGCTGGAATCCCGAAGTGCAG 2100
3521 GTGGAGATCGAGTGGGAGCTGCAGAAAGGAAACAGCAACCGCTGGAATCCCGAAATTCAG 3580
2101 TACACATCAATATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACATGCACTT 2160
3581 TACACTTCCAACTACAAACAGTCTGTAAATGTGGACTTTTACTGTGGACAAATAGCGGTG 3640
2161 TATACGTAGCTCCGCCCATTTGGCAGCCGTTACCTTACCGTCCCTGT 2209
3641 TATTACAGAGCTCGGCCCATTTGGCAGACAGTACCTGACTCGTAATCTGT 3689

RESULT 8
US-10-038-972A-12
; Sequence 12, Application US/10038972A
; Patent No. 6962815
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; PRIOR FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4679
; TYPE: DNA
; ORGANISM: adeno-associated virus 2
US-10-038-972A-12

Query Match 66.0%; Score 1458.6; DB 3; Length 4679;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;
1 ATGGCTGCGGATGGTATCTTCCAGATTTGGCTCGAGCAACCTCTCTGAGGCAATTCGC 60
2203 ATGGCTGCGGATGGTATCTTCCAGATTTGGCTCGAGGACACTCTCTGAAGGAATAAGA 2262

Qy	61	GAGTGTGGGACTTTGAAACCTGTGAGCGCCCGAAGCCCAAGCCAAACGACGAAAAGCAGGAC	120
Db	2263	CAGTGTGGGAAGCTCAAACTGTGCCCAACCAACCAAAAGCCCGACGAGCGCATAAAGGAC	2322
Qy	121	GACGGCGGGTCTGTGTCTTCTGTGCTACAAGTACTCTGGACCTCTCAACGCACTCGAC	180
Db	2323	GACAGAGGGGTCTTGTGTCTCTGTGGTACAAGTACTCTGGACCTCTCAACGCACTCGAC	2382
Qy	181	AAGGGGAGGCCCTCAACGCGCGGACGACGGCCCTCGAGCAGCAACAGGCGCTACGAC	240
Db	2383	AAGGGAGAGCGGTCAACGAGGACAGACGCGCGGCCCTCGAGCAGCAACAGGCTACGAC	2442
Qy	241	CAGCAGCTCAAGCGGGTGACAAATCCGTACTCTCGGTATACCAACGCGCAGCGCGAGTTT	300
Db	2443	CGGAGCTCGACAGCGGAGACAACCCGTACTCTCAAGTACAACACGCGCAGCGGAGTTT	2502
Qy	301	CAGGAGGCTCGACAGAGATACGTCTTTTGGGGGCAACCTCGGGGAGAGACTCTTCCAG	360
Db	2503	CAGGAGCGCTTAAGAAGAATACGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG	2562
Qy	361	GCCAAAGAAGCGGGTCTTCGAACCTCTCGGTCTGGTGTGAGGAAGCGCTAAGACGGTCCCT	420
Db	2563	GCGAAAGAAGAGGGTCTTTGAACCTCTGGGCTCTGGTGTGAGGAACCTGTTAAGACGGTCCG	2622
Qy	421	GGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCGGCACTCGGC	480
Db	2623	GGAAAGAAAGGCCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCGGAAACCGGA	2682
Qy	481	AAGACAGCCAGACGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540
Db	2683	AAGCGGGCCAGCAGCCCTGCAAGAAAGAATTGAATTTTGGTCAGACTGGAGACGCGAC	2742
Qy	541	TCAGTCCCGATCCACAACCTCTCGGAGAACCCTCCAGCAACCCCGCTGTGTGGACCT	600
Db	2743	TCAGTACTGACCCCGAGCTCTCGACAGCCACGACGCGCCCTCTGTCTGGGAACCT	2802
Qy	601	ACTCAATTGGCTTCAGCGGTGGCGCACCAATGGCAGACAAATAACGAAGCGCGCAGCGA	660
Db	2803	AATACGATGGCTACAGGCACTGGCGACCAATAAGCAGACAATAACGAGGCGCGCAGGA	2862
Qy	661	GTGGTAAATGCTCAGGAAATTTGGCATTTGGATTCCAGTGGCTGGGCGACAGAGTCATC	720
Db	2863	GTGGGTAAATTTCTCGGAAATTTGGCATTTGGATTTCACATTTGGATGGGCGACAGAGTCATC	2922
Qy	721	ACCACGACACCGCACTCGGGCTTGCCCACTTACAATAACCACTCTACAAGCAAAATC	780
Db	2923	ACCACGACACCGCACTGGGCCCTGCCCACTTACAATAACCACTCTACAACAAATT	2982
Qy	781	TCCAGTGTCTTCAACGGGGGCGACGACAAACCACTACTCTGGCTACAGCACCCCTCGG	840
Db	2983	TCCAGCCAATCA--GGAGCTCGAACGACATCACTACTTTGGCTACAGCACCCCTTGG	3039
Qy	841	GGGTATTTTGAATTTCAACAGATTTCCAATGGCACTTTTTCACAGTGACTGGCAGCGACTC	900
Db	3040	GGGTATTTTGAATTTCAACAGATTTCCAATGGCACTTTTTCACAGTGACTGGCAGCGACTC	3099
Qy	901	ATCAACAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAA	960
Db	3100	ATCAACAACAATTTGGGATTTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTCAA	3159
Qy	961	GTCAGGAGGTTCACGACGAATGTGGCGTCAACAACCTACGCTAATAACCTTACGACGACG	1020
Db	3160	GTCAAGAGGTTCACGACGAATGACGGTACGACGACGATTTGCCAATAACCTTACGACGACG	3219
Qy	1021	GTTCAAGTCTTCTCGACTCGAGATCACAGTTTCGGTACGTCTCTGGCTCTGGGCAACG	1080
Db	3220	GTTCAAGTGTTTTACTGACTCGAGATCACAGTTCCTCGTACGTCTCTGGCTCTGGGCAACG	3279
Qy	1081	GGCTGCTCTCTCGTTTCCGGCGGACGTGTTCATGATTCGGCAATACGGCTACCTGACG	1140
Db	3280	GGATGCTCTCCGCGTTTCCAGCAGACGTCTTCAATGTGCCACAGTATGATACCTCACC	3339
Qy	1141	CTCAACAATGGCAGCCAGCGGTGGGACGTTTCATCTCTTTTACTGCTCTGGAATATTTCCCT	1200

Db	3340	CTGAACAA	CGGAGT	CAGG	CAGT	AGGAC	CGCT	CTTCA	TTTTT	CTG	CGCT	GGAGT	ACT	TTCC	3399	
Qy	1201	TCTCAGAT	GTGTGAA	ACGG	CAACA	CTTT	TACCT	TCCAG	CTTAC	ACTTT	TGAGG	AAAGT	GCCT	1260		
Db	3400	TCTCAGAT	GTGTGGT	ACCG	AAACA	CACTT	TACCT	TACGT	CACTA	CTTT	TGAGG	AAAGT	GCCT	3459		
Qy	1261	TTCCACAG	CAGCTAC	CGCG	CA	CGCCAG	CGCT	TGGAC	CCGG	CTGAT	GAAT	CTCT	CTCAT	CGAC	1320	
Db	3460	TTCCACAG	CAGCTAG	CGCT	CA	CGCCAG	CAGT	CTGGA	CCGT	CTCAT	GAAT	CTCT	CTCAT	CGAC	3519	
Qy	1321	CAATACCT	GTATTAC	TGTAA	CAGAA	CTCAAA	ATCAGT	CCGG	AAAGT	CGCC	AAAA	CAAGG	AC	1380		
Db	3520	CAGTACCT	GTATTACT	TGTAG	CAGAA	CAAA	CAC	TCCA	AGTGG	AAAC	CAAC	CGCAGT	CAAGG	3579		
Qy	1381	TTGCTGTT	TAGCCGT	GGTCT	CCAG	CTGG	CATCT	TGTT	CAGCC	CCAAAA	CTGGCT	TACCT	1440			
Db	3580	CTTCAGT	TTTTCT	CAGCC	CGGAG	CGAT	GCAT	TTCCGG	GA	CCAGT	CTTAG	AACTGGCT	TTCC	3639		
Qy	1441	GGACCTGT	TATCGG	CAGC	CGCGT	TTCTAA	AAACAAA	ACAG	CAAC	AAAC	CAAC	CAACA	T	1500		
Db	3640	GGACCTGT	TATCGG	CAGC	CGCGT	TATCA	AAACAAA	ACAG	CAAC	AAAC	CAAC	CAACA	T	3699		
Qy	1501	TTTACCT	GGACT	GGTGT	TCAAA	ATATAC	CTCAAT	GGGG	TGAAT	TCCAT	CATCA	CAACCT	1560			
Db	3700	TACTCGT	GGACT	GGAGT	ACCA	CTCAAT	GGC	GAGACT	CTCT	GGTGA	ATCCG	3759				
Qy	1561	GGCACTCT	ATGGCT	CAC	AAAG	ACG	CAAG	CAAGT	TTCT	TCCAT	GAGCGGT	CTC	1620			
Db	3760	GGCCCGC	CAATGG	CAAC	AGAC	AGATGA	AA	AGTTTT	TCT	CAG	CGCGGT	3819				
Qy	1621	ATGATTTT	TGAAA	AGAG	CGCG	GAGT	TTCAAA	CACTG	CA	TGGA	CAAT	GTCA	T	1680		
Db	3820	CTCATCTT	TGGAA	CAAG	CTCAG	AAAA	CAAA	TGTG	GACAT	TGAA	AGGT	CTAT	T	3879		
Qy	1681	ACAGACGA	AGGAAAT	TAAG	CCACT	TAACCT	GTGG	CCAC	CGAA	AGATTT	TGGG	ACCGT	1740			
Db	3880	ACAGACGA	AGGAAAT	CAG	GA	CAAC	CAAT	CCCG	TGGCT	ACCG	GAGACT	TGGT	CTGTA	3939		
Qy	1741	GCAGTCA	ATTTCC	AGAC	CGAC	CGAC	CCCT	CGCA	CGGAG	ATGTG	CATG	CTAT	T	1800		
Db	3940	TCTACCA	ACCTCC	AGAG	CGCA	CAG	CAAG	CAGCT	AC	CGCAG	ATGCA	CA	CA	CAAGC	3999	
Qy	1801	GCATTACT	GGCAT	TGGT	TGG	CAAGT	PAG	AGA	CGT	GTACT	CTG	AGG	TCC	ATTTGGGC	1860	
Db	4000	GTTCTT	CCAG	CATGGT	CTGG	CAGG	ACAG	AGAT	GTGT	ACC	TT	CAG	GGG	CCCATCTGGCA	4059	
Qy	1861	AAAATTT	CTCAC	CAGAT	TGAC	ACTTT	TCAC	CCCT	CTCT	CTTAT	TGG	CGGT	TTG	GACTC	1920	
Db	4060	AAGATTCC	ACAC	CGGAC	CGACA	TTTT	TCAC	CCCT	CTCT	CTCAT	TGGT	TGGAT	TCG	GACTT	4119	
Qy	1921	AAGAAC	CGGCTCT	CTCAG	ATCT	CATCA	AAAA	CA	CGCT	CTGCT	CTG	CGAA	TCC	CTCGCGG	1980	
Db	4120	AAACAC	CCCTCT	CCAC	AGAT	TTCT	CATC	AGAA	CA	CCCC	GTACT	CTG	CGAA	TCC	TCGACC	4179
Qy	1981	GAGTTTT	CAGCTA	CAAA	GT	TGTT	CTAT	CA	CCCA	TACT	CTCA	CAG	CA	CAAGT	2040	
Db	4180	ACCTTC	AGTGG	CAAA	GT	TGTT	CTCT	CTAT	CA	CA	CAGT	ACT	CC	ACCG	4239	
Qy	2041	GTGGAA	TTGAA	TGGG	AGCT	CGA	AA	AA	CAG	CA	AGCG	CTGG	AA	TCCG	2100	
Db	4240	GTGAGAT	TCAGT	GGG	AGCT	TCG	AG	AG	AAAA	CAG	CA	AC	CGT	TGG	4299	
Qy	2101	TACACAT	CCCAAT	TATG	CAAA	ATCTG	CCAA	CGT	TGAT	TTT	ACT	GTG	GA	CA	2160	
Db	4300	TACACTT	CCAAT	TAC	AA	AGTCTG	TAA	TGTG	GACT	TTT	ACT	GTG	GA	CA	4359	
Qy	2161	TAT	ACTG	AG	CC	TG	CC	CA	AT	TGG	CA	CCG	GT	T	2209	
Db	4360	TATTCAG	AGCT	TC	CC	CA	AT	TGG	CA	CA	CA	CT	GA	CT	4408	

US-09-770-315-3									
; Sequence 3, Application US/09770315									
; Patent No. 6429001									
; GENERAL INFORMATION:									
; APPLICANT: Chiron Corporation									
; TITLE OF INVENTION: Recombinant AAV Packaging Systems									
; FILE REFERENCE: 20263-501									
; CURRENT APPLICATION NUMBER: US/09/770,315									
; CURRENT FILING DATE: 2001-01-26									
; PRIOR APPLICATION NUMBER: US 60/178,536									
; PRIOR FILING DATE: 2000-01-26									
; NUMBER OF SEQ ID NOS: 8									
; SOFTWARE: FastSeq for Windows Version 3.0									
; SEQ ID NO 3									
; LENGTH: 7557									
; TYPE: DNA									
; ORGANISM: Unknown									
; FEATURE:									
; OTHER INFORMATION: recombinant DNA									
US-09-770-315-3									
Query Match 66.0%; Score 1458.6; DB 3; Length 7557;									
Best Local Similarity 79.1%; Pred. No. 0;									
Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;									
QY	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	60						
DB	2231	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTGAGGAATAAGA	2290						
QY	61	GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCCAAGCCCAAGCAAGAGCAGGAC	120						
DB	2291	CAGTGTGGGAAGCTCAAACTCTGCCCAACCAACCAAGCCCGAGAGCGGCATAAGGAC	2350						
QY	121	GACGGCGGGTCTGGTCTCTCTGCTACAGTACTCTGGACCTTCAACGACTCGAC	180						
DB	2351	GACAGCAGGGGTCTTGTGCTTCTTGGGTACAAGTACTCTGGACCTTCAACGACTCGAC	2410						
QY	181	AAGGGGAGCCGCTCAACGGCGGCGAGCGAGCGGCCCTCGAGCAGCAACAAGGCTTAGC	240						
DB	2411	AAGGGAGAGCCGCTCAACGAGGAGAGCGCGCGGCCCTCGAGCAGCAACAAGCCTTAGC	2470						
QY	241	CAGCAGCTCAAGCGGGTGACATTCGTACTCTGGGTATACCAACGCGGACGCGGAGTTT	300						
DB	2471	CGGCAGCTCAGACGCGGAGACAACCCGTACTCTCAAGTATCAACCAACGCGGAGGTTT	2530						
QY	301	CAGAGGCTCTGCAAGAGATAGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG	360						
DB	2531	CAGGAGCGCTTAAAGAAGATAGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG	2590						
QY	361	GCCAAGAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGTCTCT	420						
DB	2591	GCGAAAAGAGGGTCTTGAACCTCTGGGCTCTGGGTGAGGAACCTGTGAAGACGGTCCG	2650						
QY	421	GGAAGAAACGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCGGGCATCGGC	480						
DB	2651	GGAAGAAAGCGGTAGAGCAGTCTCTCTGTGGAGCCAGACTCTCTCGGGAACCGGA	2710						
QY	481	AAGACAGCCAGAGCCCGCTTAAAGAGACTCAATTTTGGTCAGACTGCGGACTCAGAG	540						
DB	2711	AAGCGGGCCAGCAGCCCTGCAAGAAAAGATTGAATTTTGGTCAGACTGGAGACGCGAC	2770						
QY	541	TCAGTCCCGCATCCACAACTCTCGGAGAACCTTCAGCAACCCCGCTGTGGGACCT	600						
DB	2771	TCAGTACTGACCCCCAGCTCTCGGACAGCCACAGACGCCCTCTGGTCTGGGAACT	2830						
QY	601	ACTCAATGGCTTCAGCGGTGGCGCACCAATGSCAGACAATAACGAAGCGCGCGACGGA	660						
DB	2831	AATACGATGGCTACAGGACGTGGCGCACCAATGSCAGACAATAACAGGCGCGCGACGGA	2890						
QY	661	GTGGTAAATGCTCAGGAAATTTGGCATTTGGCATTTCCATGGCTGGGCGGACAGATCATC	720						
DB	2891	GTGGTAAATTCCTCGGGAATTTGGCAATTTGGCAATTTCCATGGATGGGCGGACAGATCATC	2950						
QY	721	ACCACAGCACCCCGCACCTTGGGCTTGGCCACCTACAATAAACCACTCTTACAGCAAAATC	780						
DB	2951	ACCACAGCACCCGAAACCTTGGGCTTGGCCACCTACAATAAACCACTCTTACAAACAAAT	3010						
QY	781	TCCAGTGTCTCAACGGGGCGGACGACACCACTACTTCTGGCTTACAGCAGCCCTCTGG	840						
DB	3011	TCAGGCCAATCA---GGAGCCTCGAACGACAAATCACTACTTTGGCTTACAGCAGCCCTTGG	3067						
QY	841	GGGTATTTTGAATTTCAACAGATTCCACTGCCACTTTTTCACCACTGCTGGCAGCACTC	900						
DB	3068	GGGTATTTTGAATTTCAACAGATTCCACTGCCACTTTTTCACCACTGCTGGCAGCACTC	3127						
QY	901	ATCAACAACAATTTGGGGATTCGGCCCAAGAGACTCAACTTCAAACCTTCAACATCCAA	960						
DB	3128	ATCAACAACAATTTGGGGATTCGGCCCAAGAGACTCAACTTCAAACCTTCAACATCCAA	3187						
QY	961	GTCAAGGAGGTCAACGACGAATGATGCGTCAACAACCATCGCTAATAACCTTACCAAGC	1020						
DB	3188	GTCAAGGAGGTCAACGACGAATGATGCGTCAACAACCATCGCTAATAACCTTACCAAGC	3247						
QY	1021	GTTCAGAGTCTTCTCGGACTCGGAGTACAGCTTTCGTAGTCTCTCGGCTCTGGCAGCAG	1080						
DB	3248	GTTCAGAGTCTTCTCGGACTCGGAGTACAGCTTTCGTAGTCTCTCGGCTCTGGCAGCAG	3307						
QY	1081	GGTGTCTCTCTCGGCTTCCGGCGGAGCTTTCATGATTTCCGCAATACGGCTTACCTGAGC	1140						
DB	3308	GGTGTCTCTCTCGGCTTCCGGCGGAGCTTTCATGATTTCCGCAATACGGCTTACCTGAGC	3367						
QY	1141	CTCAACAATTTGGGCGGAGCTTTCATCTTTTACTGCTTGGGAAATTTTCCCT	1200						
DB	3368	CTCAACAATTTGGGCGGAGCTTTCATCTTTTACTGCTTGGGAAATTTTCCCT	3427						
QY	1201	TCTCAGATCTGAGAACGGGCAACAACTTTTACCTTCAGCTTACACCTTTTGGAGGAGTGCCT	1260						
DB	3428	TCTCAGATCTGAGAACGGGCAACAACTTTTACCTTCAGCTTACACCTTTTGGAGGAGTGCCT	3487						
QY	1261	TTCCACAGAGCTACCGGCACAGCCAGAGCTGGACCGGCTGATGATCTCTCATGAC	1320						
DB	3488	TTCCACAGAGCTACCGGCACAGCCAGAGCTGGACCGGCTGATGATCTCTCATGAC	3547						
QY	1321	CAATACCTGATTTACCTGAACAGACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC	1380						
DB	3548	CAGTACCTGATTTACCTGAACAGACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC	3607						
QY	1381	TTGCTGTTTACGCTGGTCTCCAGCTGGCATCTCTTTCAGCCCAAAACCTTGGCTACCT	1440						
DB	3608	TTTACGTTTCTCAGCCCGGAGGAGTGAATTCGGGACCAAGTCTAGGAACTGGCTTCT	3667						
QY	1441	GGACCTGTTTATCGGCGAGCGCGTCTTCTAAACCAAAACACAGACAAACAGCAAT	1500						
DB	3668	GGACCTGTTTATCGGCGAGCGCGTCTTCTAAACCAAAACACAGACAAACAGCAAT	3727						
QY	1501	TTTACCTGAGCTGGTCTTCAAAATTAACCTCAATGGGGCTGAATCCATCATCAACCT	1560						
DB	3728	TACTCTGGACTGGAGCTTACCAAGTACCCTCAATGGGAGAGACTCTCTGGTGAATCCG	3787						
QY	1561	GGCACTGCTATGSCCTCACAAAGAGCAGCAAGACAAGTCTTCCATGAGCGGTGC	1620						
DB	3788	GGCCCGCCATGGCAAGCACAAGGACGATGAAGAAAAGTTTTTCTCAGAGCGGGGT	3847						
QY	1621	ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGSCATTTGGACAATGTATGGA	1680						
DB	3848	CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGACATTTGAAAAGGTCATGAT	3907						
QY	1681	ACAGACGAAGAGAAATTAAGCCACTAACCTTGTGGCCACCCGAAAGATTTGGGACCGTG	1740						
DB	3908	ACAGACGAAGAGAAATTAAGCCACTAACCTTGTGGCCACCCGAAAGATTTGGGACCGTG	3967						
QY	1741	GCAGTCAATTTCCAGAGCAGCAGACAGCCCTGCGACCGGAGATGTGCAATGTATGGGA	1800						
DB	3968	TCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCCAGATGTCAACACACAGGC	4027						
QY	1801	GCATTACCTGGCATGGTGTGGCAAGATAGAGAGCTGTACCTGCGAGGCTCCCATTTGGGCC	1860						

Db 4028 GTTCTTCCAGGCAATGGTCTGGCAGACAGAGATGTGTACTTTCAGGGGCCCATCTGGGCA 4087
Qy 1861 AAAATTCCTTCACACAGATGGACATTTTCAACCGTCTCTCTTATATGGGCGGCTTTGGACTC 1920
Db 4088 AAGATTCCACACACAGCAGACATTTTCAACCCCTCTCCCTCATGGGTGGATTGGACTT 4147
Qy 1921 AAGAACCCGCTCCCTCAGATCTCATCAAAACACGCTGTTCCTGCGAATCCTCCGGG 1980
Db 4148 AAACACCTCTCCACAGATCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACC 4207
Qy 1981 GAGTTTTTCAGCTACAAAGTTTGTCTCATTCATCACCCCAATACTCCACAGGACAAAGTGAGT 2040
Db 4208 ACCTTCAGTGGCGGCAAGTTTGTCTCTTCATCACACAGTACTCCACGGGACAGTTCAGC 4267
Qy 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCGAG 2100
Db 4268 GTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAACGCTGGAATCCCGAAATTCAG 4327
Qy 2101 TACACATCCAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTT 2160
Db 4328 TACACTTCCAATCTACAAAGTCTGTAAATGTGGACTTTTACTGTGGACACTAATGGCGTG 4387
Qy 2161 TATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTACCCGTCCTCGT 2209
Db 4388 TATTCAGAGCCTCGCCCCATTGGCACCATACCTGACTCGTAATCTGT 4436

RESULT 10

US-09-770-315-2
; Sequence 2, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 8698
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-2

Query Match 66.0%; Score 1458.6; DB 3; Length 8698;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;
Qy 1 ATGGCTCGCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTGCG 60
Db 2203 ATGGCTCGCGATGGTTATCTTCAGATTGGCTCGAGGACACTCTCTCTGAAGAAATAAG 2262
Qy 61 GAGTGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCCCAACAGCAAAAGCAGGAC 120
Db 2263 CAGTGTGGAAGCTCAAACTGGCCCAACCAACCAAGCCCGCAGCGGCAATAGGAC 2322
Qy 121 GACGGCGGGGTCTGGTCTCTCGGTACAAGTACTCTGGACCTTCAACGGACTCGAC 180
Db 2323 GACAGAGGGGTCTTGTCTCTCTGGGTACAAGTACTCTGGACCTTCAACGGACTCGAC 2382
Qy 181 AAGGGGAGCCGTCAACGGCGGAGCGAGCGGCTCTCGAGCACCAAGGCGCTACGAC 240
Db 2383 AAGGGAGCCGTCAACGAGGAGCAGCGCGGCGCTCTCGAGCACCAAGCCTACGAC 2442
Qy 241 CAGCAGCTCAAGCGGGTGACAAATCGGTATACCTCGGTATAACCAACCGCGAGTTT 300
Db 2443 CGGCAGCTCGACAGCGGAGACAAACCGTACCTCAAGTACAACCAACCGCGGAGTTT 2502

Qy 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCTGGGCGAGCAGTCTTCCAG 360
Db 2503 CAGGAGCGCTTTAAGAAAGATACGTCTTTTGGGGGCAACCTCTGGGCGAGCAGTCTTCCAG 2562
Qy 361 GCCAAGAAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAAGACGGCTCT 420
Db 2563 GCGAAAGAGAGGGTTCTTGAACCTCTGGGCTGGTTGAGGAACCTGTGTAAGACGGCTCG 2622
Qy 421 GGAAGAAACCTCGGTAGAGCAGTGGCCACAAGAGCCAGACTCTCTCTCGGCGCATTCGGC 480
Db 2623 GGAAGAAAGAGCGCGGTAGAGCAGTCTCTCTGTGAGGCCAGACTCTCTCTCTCGGGAACCGGA 2682
Qy 481 AAGCAGAGCCAGCAGCGCCGCTAAAGAGACTCAATTTTGTTCAGACTGGGAGCTCAGAG 540
Db 2683 AAGCGGCGCCAGCAGCGCTGCAAGAAAGATTTGAATTTTGTTCAGACTGGAGACGACAG 2742
Qy 541 TCAGTCCCGGATCCACAACTCTCGGAGAACTCTCCAGCAACCCCGCTGTGTGGGACCT 600
Db 2743 TCAGTACCTGACCCCGAGCTCTCGGACAGCACCGAGCGCCCTCTGTCTGGGAAT 2802
Qy 601 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATTAACGAAGGCGCCGACGGA 660
Db 2803 AATACGATGGCTACAGGCGAGTGGCGCAACAATGGCAGACAATTAACGAAGGCGCCGACGGA 2862
Qy 661 GTGGGTAAATGCTCAGGAATTTGGCATTTGCGATTCACATGGCTGGGCGCAGAGTCAATC 720
Db 2863 GTGGGTAAATCTCTCGGGAATTTGGCATTTGCGATTTCCATGGATGGGCGCAGAGTCAATC 2922
Qy 721 ACCACGAGCAGCCGACCTGGGCGCTTGGCCACCTACAAATAACCACTCTCAAGCAAAATC 780
Db 2923 ACCACGAGCAGCCGGAACCTGGGCGCTGCGCCACTACAAACCACTCTCAAAACAAAT 2982
Qy 781 TCCAGTGTCTCAAACGGGCGCAGCAACGACACCACTACTTGTGGCTACAGCACCCCTTGG 840
Db 2983 TCCAGCCCAATCA--GGAGCCTCGAAACGACAATCACTACTTGTGGCTACAGCACCCCTTGG 3039
Qy 841 GGGTATTTTGTATTTCAACAGATTCCACTGCGCACTTTTCCACCGCTGACTGGCAGGACTC 900
Db 3040 GGGTATTTTGTACTTCAACAGATTCCACTGCGCACTTTTCCACCGCTGACTGGCAAGACTC 3099
Qy 901 ATCAACAAACAATTTGGGGATTCGGGCGGACGTTTTCATGATTCGCAATAACGCTACCTCA 960
Db 3100 ATCAACAAACAATTTGGGGATTCGGGCGGACGTTTTCATGATTCGCAATAACGCTACCTCA 3159
Qy 961 GTCAAGGAGGTCAAGCAATGATGGCGTCAACCACTCGCTTAATAACCTTACAGAGCAG 1020
Db 3160 GTCAAGGAGGTCAAGCAATGATGGCGTCAAGCACTGCGATTTGCCAATAACCTTACAGAGCAG 3219
Qy 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTCTCGGCTCTGCGCAGCAG 1080
Db 3220 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTCTCGGCTCTGCGGCGCATCAA 3279
Qy 1081 GGCTGCCCTCCCTCCGTTCCCGGCGACGTTTTCATGATTCGCAATAACGCTACCTCACTGAGC 1140
Db 3280 GGAATGCTTCCGCGCTTCCAGCAGACGTTTCATGTTGGTGCACAGTATGATATCTCACC 3339
Qy 1141 CTCAACAATGGCAGCAAGCGGTGGAGCTTCATCTCTTTTACTGCTCTGGAATATTTCCCT 1200
Db 3340 CTGAACAACGGGAGTCAGGCACTAGGAGCTCTTCAATTTTACTGCTCTGAGTACTTCTCT 3399
Qy 1201 TCTCAGATGCTGAGAAAGGGCAACCTTTTACCTTCAGTACACCTTTTGGAGAGTGCCT 1260
Db 3400 TCTCAGATGCTGCGTACCGGAACAACTTTTACCTTCAGTACACCTTTTGGAGAGTGCCT 3459
Qy 1261 TTTCCACAGCAGCTAGCGGCACAGCAGCGCTTGGACCGGCTGATGAATCTCTCATCGAC 1320
Db 3460 TTTCCACAGCAGCTAGCGTCACAGCAGAGTCTGGACCGTCTCATGAATCTCTCATCGAC 3519
Qy 1321 CAATACCTGTATTA CTTGAA CAGAA CTCAAAATCAGTCTGGAGGTGCCCAAAACAGGAC 1380
Db 3520 CAGTACCTGTATTA CTTGAGCAGAA CAAACACTCCAAAGTGGAAACCAACCAACGAGTCAAGG 3579

QY 901 ATCAACAACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAA 960
Db 3030 ATCAACAACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAACTCTTAAATCAA 3089
QY 961 GTCAAGGAGGTCAACGACGAATGATGGCGTCACAAACCATCGCTTAATAACCTTACACGACG 1020
Db 3090 GTCAAGGAGGTCAACGACGAATGATGGCGTCACAAACCATCGCTTAATAACCTTACACGACG 3149
QY 1021 GTTCAAGTCTTCCGACTCGAGATACAGACTTCGATAGCTTCCGTCTCGGTCTCGGACGACG 1080
Db 3150 GTTCAAGTCTTCCGACTCGAGATACAGACTTCGATAGCTTCCGTCTCGGTCTCGGACGACG 3209
QY 1081 GGCTGCTCCCTCCGTTCCGGCGGAGCGTTTCATGATTCGCAATACGGCTACCTGACG 1140
Db 3210 GGCTGCTCCCTCCGTTCCGGCGGAGCGTTTCATGATTCGCAATACGGCTACCTGACG 3269
QY 1141 CTCAACAATGGGAGCCCAAGCCGTGGGACGTTTCACTCTTTTACTGCTGGAATATTTCCCT 1200
Db 3270 CTGACAAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAGTACTTCCCT 3329
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACTCCTCAGCTACACCTTTGAGGAAGTGGCT 1260
Db 3330 TCGCAGATGCTAAGGACTGGAAATAACTTCCAAATTCAGCTATACCTTTCGAGGATGTACT 3389
QY 1261 TTCCACAGCAGTACGCGACAGCAGCAGCCTGGACCGGCTGATGAATCCTCTCATCGAC 1320
Db 3390 TTTCACAGCAGTACGCTCACAGCAGAGTTTGGATCGGCTTATGAATCCTCTTATGAT 3449
QY 1321 CAATACCTGTATTTACCTGAACAGAACTCA---AAATCAGTCGGGAGTCCCAAAACAAG 1377
Db 3450 CAGTATCTGTACTCTGAACAGAGCAAGGACAACTCTTGAACACCAACCAATCA 3509
QY 1378 GACTGCTGTTTAGCGTGGGTCTCCAGCTGGGATGTCTGTTAGCCCAAAACTGGCTA 1437
Db 3510 CGGCTGCTTTTAGCGGCTGGGCTCAGTCTATGCTTTTGCAGGCCAGAAATTTGGCTA 3569
QY 1438 CTGGACCTGTATTCGGCAGCAGCGGTTCTTAAACAAAACAGACACACACACAGC 1497
Db 3570 CTGGGCTGCTTACCGCAACAGAGACTTCAAGAGCTGCTAAACAGACCAACACAGT 3629
QY 1498 AATTTTACCTGGACTGGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAAC 1557
Db 3630 AACTTTCTTGGACAGCGCCAGCAATATCATCTCAATGGCGGACTCGTGGTGAAT 3689
QY 1558 CTGGCAGTCTATGCTCTCACAAAGACGACGACAAAGTCTTTTCCCATGAGCGGT 1617
Db 3690 CCAGGACCACTATGCGCAGTCACAAAGGACGATGAAGAAAAATTTTCCCTATGCAACGC 3749
QY 1618 GTCATGATTTTGGAAAGAGAGCGCGGAGCTTCAACACACTGCATTGGACATGTCATG 1677
Db 3750 AATCTAATATTTGGCAAGAGGGAACAACGCAAGTAAACGCAATATGATATGTAATG 3809
QY 1678 ATTACAGACGAAGAGAAATTAAGCCACTTAACCTCTGGCCACCGAAAGATTGGGACC 1737
Db 3810 ATTACGATGAAGAGAGATTGTGACCACTATCTGTGGCAACAGCAGTATGAACT 3869
QY 1738 GTGGCAGTCAATTTCCAGAGCAGCAGCAGACAGCCCTTGCACCGGAGATGTGATGCTATG 1797
Db 3870 GTGGCAATAAATTCGAGAGCTCAATACAGCTCCACGACTGGAACGTGTCAATCATCAG 3929
QY 1798 GGAGCATTAACCTGGGATGGTGGCAAGATAGAGAGCTGTACTGCGAGGTCCTCATTTGG 1857
Db 3930 GGGGCTTAACTTGGCATGGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGG 3989
QY 1858 GCCAAAATTCCTCACACAGATGACACTTTTCAACCCGCTCTCTTATGGCGGCTTTGGA 1917
Db 3990 GCAAGATTTCTCACACGATGACACTTTTCACTCTCTCTCTCTGTATGGAGGCTTTGA 4049
QY 1918 CTCAGAAACCCGCTCTCAGATCTCTCATCAAAAAACACGCTGTTCTCGGAAATCTCCG 1977
Db 4050 CTGAACATCCGCTCTCTCAATCATGATCAAAAAATCTCCGCTAGCTCGGAATCTCTCG 4109
QY 1978 GCGGAGTTTTCAGCTACAAAGTTTGCTTCAATTCATCAACCAATACTCCACAGGACAGTG 2037

Db 4110 ACCACCTTCAGTCGGGCAAGAGTTTGTCTTCTCATCACAGTACTCCACGGACAGTCTC 4169
QY 2038 AGTGTGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTG 2097
Db 4170 AGCGTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAACGCTGGAATCCCGAAAT 4229
QY 2098 CAGTACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACATGGA 2157
Db 4230 CAGTACACATTTCAACTCAACAAGTCTGTTAATCGTGGACTTACCGTGGACTAATAGC 4289
QY 2158 CTTTATATCTAGCCTCGCCCAATTTGGCACCCGTTTACCTTACCGTCCCTGT 2209
Db 4230 GTGATTCAGAGCCTCGCCCAATTTGGCACCAAGATACCTGACTCGTAATCTGT 4341

RESULT 12
US-08-254-358-1
; Sequence 1, Application US/08254358
; Patent No. 5658785
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,358
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5658785and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-254-358-1

Query Match 65.0%; Score 1436.4; DB 2; Length 4680;
Best Local Similarity 78.7%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 466; Indels 4; Gaps 2;
QY 1 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACCAACTCTCTGAGGGCATTCGC 60
Db 2203 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAGGAATAAGA 2262
QY 61 GAGTGTGGGACTTCAAACTCGAGCCCGGAGCCCAAGCCCAACAGCAAGAGGAGGAC 120
Db 2263 CAGTGTGGAGACTCAAACTCGGCCCAACCAACCAAGCCCGAGAGCGGATAGGAC 2322
QY 121 GACGCGCGGGGTCTGGTGTCTTCTGGCTACAGTACCTCGGACCCCTTCAACGGACTCGAC 180
Db 2323 GACAGCGGGGTCTTGTGCTTCTTCTGGGTACAGTACCTCGGACCCCTTCAACGGACTCGAC 2382

Qy	181	AAGGGGAGCCCGTCAACGCGGGAGCGCAGCGGCCTTCGAGCAGCAAGGCTTACGAC	240
Db	2383	AAGGGAGAGCCGGTCAACGAGGCGAGCGCGCGGCCCTCGAGCAGCAAAAGCTTACGAC	2442
Qy	241	CACGAGCTCAAGCGGTGACNATCCGTACTCGGTATATACCAACGCCGACGCGGAGTTT	300
Db	2443	CGGCGAGCTCGACAGCGGAGACAACCCGTACTCAAGTACAAACACGCGGACGCGGAGTTT	2502
Qy	301	CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTTCGGGCGAGCAGTCTTCCAG	360
Db	2503	CAGGAGCGCTTAAAGAAGATACGTCTTTTGGGGGCAACCTTCGGAACGAGCAGTCTTCCAG	2562
Qy	361	GCCAAGAAGCGGGTTCTCGAACTCTCGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT	420
Db	2563	GCGAAAAGAGGGTCTTGAACCTCTCGGCTCGGTTGAGGAACCTGTTAAGACGGCTCCG	2622
Qy	421	GGAAGAAACGTCGCGGTAGAGCAGTGCSCACAGAGCCAGACTCTCTCGGCACTGGC	480
Db	2623	GGAAAAAGAGCGGTAGAGCAGTCTCTGTGGAGCCAGACTCTCTCGGGAACCGGA	2682
Qy	481	AAGACAGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540
Db	2683	AAGCGGGCCAGCAGCCTGCAAGAAAPAGATTGAAATTTTGGTCAGACTGGAGACGAGAC	2742
Qy	541	TCAGTCCCGANTCAAACTCTCGGAGAACTCTCAGCAACCCCGCTGCTGTGGACCT	600
Db	2743	TCAGTACTCGACCCCAAGCCTCTCGGACAGCCACAGCAGCCCGCTCTGGTCTGGAACT	2802
Qy	601	ACTACANTGCTTCAGCGGTGGCGCACCAATGSCAGACAATAACGAAGCGCGGACGGA	660
Db	2803	ANTACGATGGCTACGCGAGTGGCGCACCAATGSCAGACAATAACGAAGCGCGGACGGA	2862
Qy	661	GTGGGTAAATGCCTCAGGAAATTTGGCAATTGCGATTTCCAATGGCTGGGCGCAGAGATCATC	720
Db	2863	GTGGGTAAATCTCTCCGGAATTTGGCATTTGCGATTTCCACATGGATGGGCGACAGATCATC	2922
Qy	721	ACNACGACGCCGACCTGGGCCCTTGGCCACTGACATTAACNACCTCTTACAGCAATC	780
Db	2923	ACNACGACGCCGAACTGGGCCCTGCCCACTTACAAACCAACCACTCTTACAAACAAAT	2982
Qy	781	TCCAGTGCTTCAACGGGGGCGAGCAACGACAACCACTACTTTGGCTACAGCACCCCTGG	840
Db	2983	TCCAGCCAAATCA---GAGAGCTCGAAGCAATCACTACTTTGGCTACAGCACCCCTGG	3039
Qy	841	GGGTATTTTGATTTCAACAGATTTCCACTGGCCACTTTTCAACAAGTGAATGCGACGACTC	900
Db	3040	GGGTATTTTGACTTTCAACAGATTTCCAATGGCCACTTTTCAACAAGTGAATGCGACGACTC	3099
Qy	901	ATCAACAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTTTTCAACATCCAA	960
Db	3100	ATCAACAACAACTGGGGATTTCCGACCCCAAGAGACTCAACTTTTCAAGCTCTTTTAAACATTCAA	3159
Qy	961	GTCAAGGAGGTACGACGAATGATGGCGTCAACACCATCGCTATAACCTTTACGACGACG	1020
Db	3160	GTCAAGAGGTACGACGAATGACGGTACGACGACGATTTGCCAATAACCTTTACGACGACG	3219
Qy	1021	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAAGTCTCTCGGCTCTGGGACCAAG	1080
Db	3220	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAAGTCTCTCGGCTCTGGGACCAAG	3279
Qy	1081	GGCTGCCTCCCTCGGTTCCCGGCGAGTGTTCATGATTCGGCAATACGGCTACCTGACG	1140
Db	3280	GGATGCCTTCCCGGTTTCCCAAGCAGACGTCTTCAATGTTGCCAATGATGATACCTCACC	3339
Qy	1141	CTCAACAATGGCAGCAAGCCGTGGGAGCTTTCATCTTTTACTGCCCTGGGAATATTTCCCT	1200
Db	3340	CTGAACAACGGAGTCAAGCAGTAGGAGCGCTCTTCAATTTTACTGCCCTGGAGTACTTTCT	3399
Qy	1201	TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT	1260
Db	3400	TCTCAGATGCTGCTTACGGAACAACTTTACCTTCAGCTACACTTTTGGAGGAGCTTCT	3459

QY	1361	TTCCACAGCAGCTACGGCGCACAGCCAGAGCGCTGGACCGGCTGATGATTCCTCTCATCGAC	1322
DB	3460	TTCCACAGCAGCTACGCTCA CAGCCAGAGTCTGGACCGTCTCATGAATTCCTCTCATCGAC	3519
QY	1321	CAATACCTGTATTACCTCTGAACAGAACTCAAAATCAGTCCGGAGTGCCTCAAAACAAGGAC	1380
DB	3520	CAGTACCTGTATTACTTTGAGCAGNACAAACACTCCAAGTGGAAACCAACCGCAGTCAAGG	3579
QY	1381	TTGCTGTTTAGCCGTGGGTCTTCAGCTGGCATGTCTGTTCAGGCCAAAAAATCGGTACCT	1440
DB	3580	CTTCAGTTTTCTCAGCGCGGAGCGATGACATTCGGGACCAAGTCTAGGAACTGGCTTCCT	3639
QY	1441	GGACCCCTGTTATCGGCGAGCAGCGGTTTCTAAACAAAAACAGACAAACAACAGCAAT	1500
DB	3640	GGACCCCTGTTATCCGCCACGACGCGAGTATCAAGACATCTCGGATATACAAACAACAGTGAA	3699
QY	1501	TTTTACCTGGACTGGTGCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCCCT	1560
DB	3700	TACTCTGGTGGACTGGAGCTACCAAGTACCACTCAATGGCAGAGACTCTCTGGTGAATCCG	3759
QY	1561	GGCACTGCTATGGCCTCACACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGC	1620
DB	3760	GGGCCCGCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCAGAGCGGGGTT	3819
QY	1621	ATGATTTTTGAAAAAGAGAGCGCCGAGCTTCAACACTGCATGGATGGACAATATCATGATT	1680
DB	3820	CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGAACATGTGAAGAGGTCATGATT	3879
QY	1681	ACAGACGAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGTG	1740
DB	3880	ACAGACGAAGAGGAAATTCGAACAACCAANTCCGCTGGCTACGGAGCAGTATGGTCTTGTA	3939
QY	1741	GCAGTCAAAATTTCCAGAGCAGCAGCACAGACCCCTGCGACCGGAGATGTGCATGTATGGGA	1800
DB	3940	TCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAAGGC	3999
QY	1801	GCAATTAAGCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTCTCAGGGTCCCATTTGGGCC	1860
DB	4000	GTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTTCAGGGGGCCCATCTGGGCA	4059
QY	1861	AAAATTCCTCACAGATGGACACTTTTCAACCGTCTCCTCTTATGGGCGGCTTTGGACTC	1920
DB	4060	AAGATTCACACACGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCGGACTT	4119
QY	1921	AAGAACCCGCTCCTTCAGATCCTCATCAAAAAACAGCGCTGTTCCTGCGGAATTCCTCGCGG	1980
DB	4120	AAACACCCCTCTCCACAGATTCTCATCAAGNACACCCCGGTACCTCGGAATCCTTCGACC	4179
QY	1981	GAGTTTTACGCTACAAAGTTTGGTTCATTATCATCAACCAATCTATCCACAGGACA-AGTGAG	2039
DB	4180	ACCTTCAGTGGCGCAAGTTTGTCTTCTCTCATCACACAGTACTCCACGGGACACGGTCAG	4239
QY	2040	TGTGGAATTTGAATTTGGGAGCTGCAGAAAGAAAACAGCAGCGCTGGAAATCCCGAAGTGCA	2099
DB	4240	CGTGGAGATCGAGTGGGAGCTGTCAGNAGGAAAACAGCAAAACGCTTGGAAATCCCGAAAATTC	4299
QY	2100	GTACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACT	2159
DB	4300	GTACACTTCAAATCAACAAGTCTGTAACTCGTGGACTTACCGTGGTACTAATGCGGT	4359
QY	2160	TTATATGAGCCTCGCCCCATTTGGCACCCGTTACTTTACCGGTCCCCGTG	2209
DB	4360	GTATTCAGAGCTCTGGCCCATTTGGCACCCAGATACCTGACTGTAATCTGTG	4409

RESULT 13
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus
; TITLE OF INVENTION: Methods of Gene
; Delivery

1 NUMBER OF SEQUENCES: 3
2 CORRESPONDENCE ADDRESS:
3 ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
4 STREET: 6300 Sears Tower, 233 S. Wacker Drive
5 CITY: Chicago
6 STATE: Illinois
7 COUNTRY: USA
8 ZIP: 60606
9
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Floppy disk
12 COMPUTER: IBM PC compatible
13 OPERATING SYSTEM: PC-DOS/MS-DOS
14 SOFTWARE: PatentIn Release #1.0, Version #1.25
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/475,391
17 FILING DATE: 07-JUN-1995
18 CLASSIFICATION: 435
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: 08/254,358
21 FILING DATE:
22 ATTORNEY/AGENT INFORMATION:
23 NAME: No. 578621land, Greta S.
24 REGISTRATION NUMBER: 35,302
25 REFERENCE/DOCKET NUMBER: 31975
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (312) 474-6300
28 TELEFAX: (312) 474-0448
29
30 INFORMATION FOR SEQ ID NO: 1:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 4680 base pairs
33 TYPE: nucleic acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: DNA (genomic)
37
38 US-08-475-391-1

Query Match 65.0%; Score 1436.4; DB 2; Length 4680;
Best Local Similarity 78.7%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 466; Indels 4; Gaps 2;

QY 1 ATGGCTGCCGATGGTTATCTCCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTGCG 60
DB 2203 ATGGCTGCCGATGGTTATCTCCAGATTGGCTCGAGGACAACTCTCTGAGGAAATAGA 2262
QY 61 GAGTGTGGGACTTGAACCTTGAGCCCGGAGCCCAAGCCCAAGCCCAAGCAAGAGGAG 120
DB 2263 GAGTGTGGGAGCTCAACCTTGCCGCCACCACCACCAAGCCCGCAGCGCATAGGAC 2322
QY 121 GACGGCCGGGCTGTGGTCTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC 180
DB 2323 GACAGAGGGGCTTGTGCTTCTCGGTTCAAGTACCTCGGACCTTCAACGGACTCGAC 2382
QY 181 AAGGGGAGCCGTCAACGCGGCGGAGCGAGCGGCCCTCGAGCAGCAGAGGCGCTACGAC 240
DB 2383 AAGGGAGCCGCTCAACGAGGCGAGCGCGCGGCCCTCGAGCAGCAGAGGCGCTACGAC 2442
QY 241 CAGCAGCTCAAGCGGGGTGACAACTCGGTATACCTCGGTATAAACACCGCAGCGAGTTT 300
DB 2443 CGGCAGCTCGACAGCGGAGACACCGGTACCTCAAGTACAAACACCGCAGCGAGTTT 2502
QY 301 CAGGAGGCTCGAAGAGATAGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360
DB 2503 CAGGAGCGCTTAAAGAAGATAGTCTTTTGGGGGCAACCTCGGAGCAGCAGTCTTCCAG 2562
QY 361 GCCAAGAAGCGGTTCTCGAACCTCTCGGTCTGTTGAGGAGCGCTAAGCGGCTCCT 420
DB 2563 GCGAAAAGAGGGTCTTGAACCTCTGCGCTGTTGAGGAACTCTGTTAAGACGGCTCCG 2622
QY 421 GGAAGAAAGCGTCCGGTAGAGCAGTGCACCAAGAGCCAGACTCTCTCGGCGATCGGC 480
DB 2623 GGAAGAAAGCGCGGTAGAGCACTCTCTGTGGAGCAGACTCTCTCTCGGAAACCGGA 2682

QY 481 AAGACAGCCAGCAGCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGGACTCAGAG 540
DB 2683 AAGCGGCCAGCAGCCCTGCAAGAAAAGATTGAAATTTTGGTCAGACTGGAGACGAGAC 2742
QY 541 TCAGTCCCGGATCCACAACCTCTCGGAGAACTCCAGCAACCCCGCTGTGTGGGACT 600
DB 2743 TCAGTACCTGACCCCGAGCTCTCGGACAGCCACCGAGCCCTCTGTGTGGGAACT 2802
QY 601 ACTACAATGGCTTCAGGCGGTGGCGACCAATGGCAGACAAATAAGAAAGGCGCGACGGA 660
DB 2803 AATAGATGGCTACAGGCAAGTGGCGACCAATGGCAGACAAATAAGAAAGGCGCGACGGA 2862
QY 661 GTGGTAATGCTCAGGAAATTTGGCAATTTGCCATTTCCACATGGCTGGGCGAGAGTCAATC 720
DB 2863 GTGGTAATTTCTCCGAAATTTGGCAATTTGCCATTTCCACATGGATGGGCGAGAGTCAATC 2922
QY 721 ACCACAGCAGCCCGCAGCTGGGCTTGGCCACCTACATAATACCACTCTACAAGCAAAATC 780
DB 2923 ACCACAGCAGCCCGCAACCTGGGCTTGGCCACCTACATAATACCACTCTACAAGCAAAATC 2982
QY 781 TCCAGTGTCTCAACGGGGGCGAGCAACGACAACTACTCTTCGGCTCAGACACCCCTGG 840
DB 2983 TCCAGCCAAATCA--GGAGCCTCGAAACGACAACTACTACTTTTGGCTCAGACACCCCTGG 3039
QY 841 GGGTATTTTGAATTTCAACAGATTCCACTGCCACTTTTCCACACGTCGACTGGCAGCGACTC 900
DB 3040 GGGTATTTTGAATTTCAACAGATTCCACTGCCACTTTTCCACACGTCGACTGGCAGCGACTC 3099
QY 901 ATCAACAACTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCAA 960
DB 3100 ATCAACAACTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAGCTCTTTAAACATCAA 3159
QY 961 GTCAAGAGGTCACGCAATGATGGCTCAACCACTCGCTATAATACCTTACCGAGCAG 1020
DB 3160 GTCAAGAGGTCACGCAATGATGGCTCAACCACTCGCTATAATACCTTACCGAGCAG 3219
QY 1021 GTTCAAGTCTTCTCGGACTCCGAGTACCACTTCCGTCAGCTTCCGCTCTGCGCAGCAG 1080
DB 3220 GTTCAAGTCTTCTCGGACTCCGAGTACCACTTCCGTCAGCTTCCGCTCTGCGCAGCAG 3279
QY 1081 GGTCTCCCTCCCTCCCTCCCGCGGACGCTGTTCAATGATTCGCGCAATACGGCTACCTGACG 1140
DB 3280 GGTCTCCCTCCCGCTTCCAGCAGACGCTTTCATGGTGCCACAGTATGGATACCTCACC 3339
QY 1141 CTCACAATGGCAGCAGCAGCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCT 1200
DB 3340 CTCACAATGGCAGCAGCAGCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCT 3399
QY 1201 TCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGTACACCTTTTGGAGAGTGCCT 1260
DB 3400 TCTCAGATGCTGCTGACCGGAAACAACTTTTACCTTTCAGTACACCTTTTGGAGAGTGCCT 3459
QY 1261 TTCCACAGCAGCTACCGCAGCAGCAGCGCTGGAACCGGCTGATGAATCTCTCATCGAC 1320
DB 3460 TTCCACAGCAGCTACCGCTCAGCAGCAGAGTCTGGACCGCTCTCATGAATCTCTCATCGAC 3519
QY 1321 CAATACCTGTATTAACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAGGAC 1380
DB 3520 CAGTACCTGTATTAACCTGAGCAGAAACAACTTCAAGTGGAAACCAACCAACGAGTCAAGG 3579
QY 1381 TTGCTGTTTACCGGTGCTTCCAGCTGGCATGCTGTTTACGCTTCCAGCCCAAACTCTGCTACCT 1440
DB 3580 TTGCTGTTTACCGGTGCTTCCAGCTGGCATGCTGTTTACGCTTCCAGCCCAAACTCTGCTACCT 3639
QY 1441 GGACCTCTGTTATTCGCGCAGCGCGCTTTCTTAAACAAACAAACAGCAACCAACAGCAAT 1500
DB 3640 GGACCTCTGTTATTCGCGCAGCGCGAGTATCAAGAGCATCTCGCGATTAACAAACAGTGA 3699
QY 1501 TTATCTGAGCTGTGCTTCAAAATATTAACCTCAATGGGCGGTGAATCTCATCAACCT 1560
DB 3700 TACTGTGAGCTGTGAGTACCAAGTACCACTCAATGGCAGAGACTCTCTGTGTGAATCCG 3759
QY 1561 GGCAGCTGTATGGCTTCACACAAAGCAGGAAAGCAAGTCTTCTTCCCATGAGCGGTGC 1620

Db 3760 GGCGCCGCGCATGGCAAGCCACAGGACGATGAAGAAAAGTTTTTCTCCAGAGCGGGTT 3819
Qy 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAACACACTGCTTGGACAAATGTCATGATT 1680
Db 3820 CTCATCTTTGGAGCAGAGCTCAGAGAAAACAAATGTGAACATTTGAAAAGTTCATGATT 3879
Qy 1681 ACAGACGAAGAGAAATTAAGCCACTAAACCTCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
Db 3880 ACAGACGAAGAGAAATCGGAACCAATCCGCTGCTAGCGAGCAGTATGTTCTGTA 3939
Qy 1741 GAGTCAATTTCCAGAGCAGCAGCAGACAGACCCCTGCGACCGGAGATGTGATGTATGGA 1800
Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAAGGC 3999
Qy 1801 GCATTAACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGCGAGGTCCCATTTGGGCC 1860
Db 4000 GTTCTCCAGGATGTGTGGCAGGACAGATGTGTACTTCCAGGGGCCCATCTGGGCA 4059
Qy 1861 AAAATTCCTCACACAGATGGACACTTTTCAACCCGCTCTCTTTATGGCGGCTTTGGACTC 1920
Db 4060 AGATTTCCACACAGCAGCGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCGGACTT 4119
Qy 1921 AGAACCCGCTCTCAGATCTCTATCAAAACACACGCTGTCTCTCGGAATCTCCGGCG 1980
Db 4120 AAACACCTCTCTCCACAGATTTCTCATCAAGAACACACCCCGGTACTCTCGAATCTCTCGACC 4179
Qy 1981 GAGTTTTCAGCTACAAGTTTGTCTTCAATCATCACCCCAATCTCCACAGGAC - AGTGAG 2039
Db 4180 ACCTTCAGTGGCGCAAGTTTGTCTCTTCATCACACAGTACTCCACGGGACACGCTCAG 4239
Qy 2040 TGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAACTGCA 2099
Db 4240 CGTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAATTC 4299
Qy 2100 GTACACATCAATATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAACATGGACT 2159
Db 4300 GTACACTTCAACTACAAACAAGTCTGTAACTGTGGACTTACCGTGGATACTAATGGCGT 4359
Qy 2160 TTATACTGAGCTCGCCCATTTGGCACCGGTACTTACCTTACCCGCTCCCTGT 2209
Db 4360 GTATTACAGGCCTCGCCCAATTTGGCACCATGACTGACTCGTAATCTGT 4409

RESULT 14

US-08-709-609-1
; Sequence 1, Application US/08709609
; Patent No. 5858775
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,609
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5858775and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975

TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-709-609-1
Query Match 65.0%; Score 1436.4; DB 2; Length 4680;
Best Local Similarity 78.7%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 466; Indels 4; Gaps 2;
Qy 1 ATGGCGTCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTCCG 60
Db 2203 ATGGCTCGCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAGGAATAAGA 2262
Qy 61 GAGTGTGGGACTTGAACCTCGAGCCCGAAAGCCCAAGCCAAAGCCAGCAAAAGCAGGAC 120
Db 2263 GAGTGTGGAAGCTCAACCTGGCCACCAACCACCAAGCCCGCAGAGCGGATAGGAC 2322
Qy 121 GACGCGCGGGTCTGGTCTTCTGGCTACAAAGTACCTCGGACCTTTCAACGGAATCGAC 180
Db 2323 GACAGCAGGGTCTTGTGCTTCTGGGTACAAAGTACCTCGGACCTTTCAACGGAATCGAC 2382
Qy 181 AAGGGGAGCCCCGTCAACGCGCGGACGCGGCGCTCGACACGACGACGAGCCCTACGAC 240
Db 2383 AAGGGAGAGCCGGTCAACGAGGACGCGCGCGGCTCGACACGACGACGACGACGAC 2442
Qy 241 CAGCAGCTCAAGCGGGTGCACAACTCCGTACCTCGGTATAACACACGCGCCGAGTTT 300
Db 2443 CGCGAGCTCGACAGCGGAGACAAACCGTACCTCAAGTACAAACACGCGCGAGGAGTTT 2502
Qy 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
Db 2503 CAGGAGCGCTTAAAGAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562
Qy 361 GCCAAGAGCGGGTCTCGAACTCTCGTCTGGTTGAGGAGGCGCTAAGAGCGCTCCT 420
Db 2563 GCGAAGAGAGGGTCTTGAACCTCTGGGCGCTGGTTGAGGAACCTGTTAAGACGGCTCG 2622
Qy 421 GGAAGAGAAAGTCCGCTAGAGCAGTCGCCAAGAGCCAGACTCTCTCTCGGCGATCGGC 480
Db 2623 GGAAGAGAGCGCGGTAGAGCAGTCTCTCTGTGGAGCCAGACTCTCTCTCGGGAACGGA 2682
Qy 481 AAGACAGGCGCAGCAGCCCTGCAAGAGAGACTCAATTTTGGTCAAGACTGGCGACTCAGAG 540
Db 2683 AAGCGGGCCAGCAGCCTGCAAGAGAAAGATTGAATTTTGGTCAAGACTGGAGACGAGAC 2742
Qy 541 TCAGTCCCCGATCCAACTCTCGGAGAACCTTCAGCAACCCCGCTCTCTGTTGGGACCT 600
Db 2743 TCAGTACCTGAGCCCCAGCCTCTCGGACAGCCACCAAGAGCCCTCTCTGTTGGGAACT 2802
Qy 601 ACTACAATGGCTTCAGGCGGTGCGCACCAATAGGACAGCAATAACGAAGCGCGCAGCGA 660
Db 2803 AATACGATGGCTACAGGAGTGGCGCAATGGCAGACATTAACGAGGCGCGCAGCGA 2862
Qy 661 GTGGGTAATTCCTCAGGAAATTTGGCAATTCGATTTCCATGTTGGTGGCGCAGAGTCAATC 720
Db 2863 GTGGGTAATTCCTCGGAAATTTGGCAATTCGATTTCCATGTTGGTGGCGCAGAGTCAATC 2922
Qy 721 ACCACGACGCGCAGCTTGGGCTTGGCCCACTCAATAAACAACCTCTTACAGCAATC 780
Db 2923 ACCACGACGCGAACCCTGGGCTTGGCCCACTCAATAAACAACCTCTTACAAACAAAT 2982
Qy 781 TCCAGTGTCTCAACGGGGCGCAGACACCACTACTTTCGGCTACAGACCCCTCTGG 840
Db 2983 TCCAGCCCAATCA---GGAGCTCTGAAACGACAACTACTTTCGGCTACAGACCCCTTGG 3039

841 GGGTATTTTGAATTTCAACAGATTTCACATGCGACTTTTCCACAGCTGAGTGGCAGGACTC 900
Db |||||
3040 GGGTATTTTGAATTTCAACAGATTTCACATGCGACTTTTCCACAGCTGAGTGGCAGGACTC 3099
Qy |||||
901 ATCAACAACAAATTTGGGGATTCGCGCCCAAGAGACTCAACTTCAAACTCTTTCAACATCCAA 960
Db |||||
3100 ATCAACAACAAATTTGGGGATTCGCGCCCAAGAGACTCAACTTCAAGCTCTTTAAATTCAA 3159
Qy |||||
961 GTCAAGGAGGTTCAGACGAGTATGCGGTTCACAACTGCTGCTTAACCTTACAGAGAGG 1020
Db |||||
3160 GTCAAGGAGGTTCAGACGAGTATGCGGTTCACAACTGCTGCTTAACCTTACAGAGAGG 3219
Qy |||||
1021 GTTCAAGTCTTCTCGAGCTCGAGTACCACTTTTCACTTCACTTCACTTCACTTCACTTCACTT 1080
Db |||||
3220 GTTCAAGTCTTCTCGAGCTCGAGTACCACTTTTCACTTCACTTCACTTCACTTCACTTCACTT 3279
Qy |||||
1081 GGTGCTCTCCCTCGGTTCGCGCGGAGAGCTTTCATGATTTCCGCAATACGCTTACCTGAGG 1140
Db |||||
3280 GGTGCTCTCCCTCGGTTCGCGCGGAGAGCTTTCATGATTTCCGCAATACGCTTACCTGAGG 3339
Qy |||||
1141 CTCACAATTTGGGAGTTCGCGCGGAGAGCTTTCATGATTTCCGCAATACGCTTACCTGAGG 1200
Db |||||
3340 CTCACAATTTGGGAGTTCGCGCGGAGAGCTTTCATGATTTCCGCAATACGCTTACCTGAGG 3399
Qy |||||
1201 TCTCAGATGCTGAGACGCGGAGAGCTTTCATGATTTCCGCAATACGCTTACCTGAGGAGT 1260
Db |||||
3400 TCTCAGATGCTGAGACGCGGAGAGCTTTCATGATTTCCGCAATACGCTTACCTGAGGAGT 3459
Qy |||||
1261 TTTCCACAGAGCTTACGCGCACAGGAGCTTTCGAGCGGCTGATGATTTCCGCAATACGCTTAC 1320
Db |||||
3460 TTTCCACAGAGCTTACGCGCACAGGAGCTTTCGAGCGGCTGATGATTTCCGCAATACGCTTAC 3519
Qy |||||
1321 CAATACCTGATTTACCTGAGAGAGCTTCAAAATTCAGTCCGAGAGTGGCCCAAAACAGAGAC 1380
Db |||||
3520 CAGTACCTGATTTACCTGAGAGAGCTTCAAAATTCAGTCCGAGAGTGGCCCAAAACAGAGAC 3579
Qy |||||
1381 TTGCTGTTTATGCGGTCTCCAGCTGCGAGTCTGTTTCCGCGCAAACTGGGTACTT 1440
Db |||||
3580 TTGCTGTTTATGCGGTCTCCAGCTGCGAGTCTGTTTCCGCGCAAACTGGGTACTT 3639
Qy |||||
1441 GGACCTGTTTATGCGGAGAGGCGGTCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 1500
Db |||||
3640 GGACCTGTTTATGCGGAGAGGCGGTCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 3699
Qy |||||
1501 TTATCTGAGTGTGCTTCAAAATTAACCTCAATGCGGCTGATTCATTCATTCATTCATTCAT 1560
Db |||||
3700 TACTCTGAGTGTGCTTCAAAATTAACCTCAATGCGGCTGATTCATTCATTCATTCATTCAT 3759
Qy |||||
1561 GGCACCTGCTATGCTTCAAAACAGAGAGCTTCAAAATTCAGTCCGAGAGTGGCCCAAAACAG 1620
Db |||||
3760 GGCACCTGCTATGCTTCAAAACAGAGAGCTTCAAAATTCAGTCCGAGAGTGGCCCAAAACAG 3819
Qy |||||
1621 ATGATTTTGGAAAGAGAGGCGCGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 1680
Db |||||
3820 ATGATTTTGGAAAGAGAGGCGCGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 3879
Qy |||||
1681 ACAGACGAGAGAGAAATTAAGCCACTAAACCTGCTGCGCCCAAGAGAGTGGGACCGGTG 1740
Db |||||
3880 ACAGACGAGAGAGAAATTAAGCCACTAAACCTGCTGCGCCCAAGAGAGTGGGACCGGTG 3939
Qy |||||
1741 GCAGTCAATTTCCAGAGAGAGAGAGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 1800
Db |||||
3940 GCAGTCAATTTCCAGAGAGAGAGAGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 3999
Qy |||||
1801 GCATTAACCTTGGATGCTGCGAGAGAGAGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCT 1860
Db |||||
4000 GTTCTTCCAGGATGCTGCGAGAGAGAGAGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCT 4059
Qy |||||
1861 AATATTTCTCAACAGAGAGAGAGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 1920
Db |||||
4060 AATATTTCTCAACAGAGAGAGAGAGCTTCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAA 4119
Qy |||||
1921 AAGAACCCGCTCTCAGATCTCTATCAAAACAGAGAGCTTCAAAACAGAGAGCTTCAAAACAGAG 1980

Db |||||
4120 AAAACACCTCTCCACAGATTCTCATCAGAACACCCCGGTACCTGCGAATCTCTCGACC 4179
Qy |||||
1981 GAGTTTTCAGCTACAAAGTTTGTCTTATTCATCACCACCAATATCTCCACAGGACA-AGTGAG 2039
Db |||||
4180 ACCTTCAGTGGCGGAAAGTTTGTCTTCTATCACAAGTACTCCACCGGACACGGTCAG 4239
Qy |||||
2040 TGTGGAATTTCAATGGGAGCTGCAGAAAGAAAAGCAAGCGCTGGAATCCGGAAGTCA 2099
Db |||||
4240 CGTGAGATTCAGTGGGAGCTGCAGAAAGAAAAGCAAGCGCTGGAATCCGGAATTC 4299
Qy |||||
2100 GTACACATCTCAATATGCAAAATCTGCCAAAGTTGATTTTACTGTGGACAAACAATGGACT 2159
Db |||||
4300 GTACACTTCAACTACAAAGTCTGTTAATCGTGGACTTACCGTGGGATCTAATGGCGT 4359
Qy |||||
2160 TTATACTGAGCTCGCCCATTTGGCACCGCTTACTTACCGTCCCTGT 2209
Db |||||
4360 GTATTTCAGAGCTCGCCCATTTGGCACAGATACCTGACTCGTAATCTGT 4409

RESULT 15

PCT-US95-07178-1

; Sequence 1, Application PC/TUS9507178

; GENERAL INFORMATION:

; APPLICANT: Johnson, Philip R.

; TITLE OF INVENTION: Adeno-Associated Virus Materials and

; TITLE OF INVENTION: Methods

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 6300 Sears Tower, 233 S. Wacker Drive

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/07178

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Noland, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 31975

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 474-6300

; TELEFAX: (312) 474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4680 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

PCT-US95-07178-1

Query Match 65.0%; Score 1436.4; DB 6; Length 4680;

Best Local Similarity 78.7%; Pred. No. 0;

Matches 1740; Conservative 0; Mismatches 466; Indels 4; Gaps 2;

Qy 1 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACCAACTCTCTGAGGGCATTCGC 60

Db 2203 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAGGAATAAGA 2262

Qy 61 GAGTGGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG 120

Db 2263 CAGTGGTGGGAGCTCAAACTGGCCCAACCCACCAACCCCAAGCCCGGAGGCGGATAAGGAC 2322

121 GACGGCCGGGTCTGTGTTCTCTGGCTACAGTACTCTGGACCTTTCACAGGACTCGAC 180
122 |||||
2323 GACAGCAGGGTCTTGTGTTCTCTGGGTACAGTACTCTGGACCTTTCACAGGACTCGAC 2382
181 AAGGGGAGCCGTCAACGGGGGAGCAGCGGCCCTCGAGCAGCAAGGCTTACGAC 240
2383 AAGGGAGAGCCGTCAACGAGGAGCAGCGCGGCCCTCGAGCAGCAAGGCTTACGAC 2442
241 CAGCAGCTCAAGCGGGTGACAACTCCGTACTCTGGGTATAACCAACCGCCGACGCGGAGTTT 300
2443 CGGCAGCTGCAGCAGCGAGCAACCCGTACTCAAGTACACCAACCGCCGAGCGGAGTTT 2502
301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG 360
2503 CAGGAGCGCTTAAAGAAATACGTCTTTTGGGGCAACCTCGGAGCAGCAGTCTTCCAG 2562
361 GCCAAGAGCGGGTCTCGAAGCTCTCGGTCTGGTTGAGGAGCGCTAAGACGGCTCCT 420
2563 GCGAAAAGAGGGTCTTGAACCTCTGGGCTTGGTTGAGGAACCTGTAAAGACGGGTCCG 2622
421 GGAAGAAAGCTCCGGTAGAGCAGTCGCCAAGAGCCAGACTCTCTCGGGCATCGGC 480
2623 GGAAAAAGAGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCGGGAACCGGA 2682
481 AAGACAGGCGCAGAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG 540
2683 AAGGCGGCGCAGCAGCTTCAAGAAAAGATTGAATTTTGGTTCAGACTGGAGACGCGAG 2742
541 TCAGTCCCCGATCCACACTCTCGGAGAACCTTCAGCAACCCCGCTGCTGTGGACCT 600
2743 TCAGTACCTGACCCCCAGCCCTCTCGAGCAGCCACCGAGCCCTCTGCTGTGGAACT 2802
601 ACTACAAATGCTTCAGCGGTGCGCACCAATGGCAGACAATAACGAAGCGCGCAGCGA 660
2803 AATACAGATGCTACAGCAGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGCGA 2862
661 GTGGGTAAATGCTTCAGGAAATGGCAATTCGCAATTCACATGGCTGGCGCAGAGTCATC 720
2863 GTGGGTAAATTCCTCCGAAATTCGCAATTCGCAATTCGCAATGGATGGCGCAGAGTCATC 2922
721 ACCACAGCAACCCGACCTGGGCTTGCACCACTTACAAATACCACTCTTACAGCAATC 780
2923 ACCACAGCAACCCGACCTGGGCTTGCACCACTTACAAATACCACTCTTACAAACAAAT 2982
781 TCAGTGTCTTACAGGGGCGCAGCAGCAACCACTTCTCGGTACAGCAACCCCTGG 840
2983 TCAGGCAATCA--GGAGCCTCGAAGCAATCACTACTTTGGCTACAGCAACCCCTGG 3039
841 GGGTATTTTGAATTTCAACAGATTCACCTGCGCACTTTTACACAGTGACTGGCAGCACTC 900
3040 GGGTATTTTGAATTTCAACAGATTCACCTGCGCACTTTTACACAGTGACTGGCAGCACTC 3099
901 ATCAACAAATTTGGGATTCGGGCTTCAAGAGACTCAACTTCAAACTTCTTCAACATCCAA 960
3100 ATCAACAAATTTGGGATTCGGGCTTCCAGCCCAAGAGACTCAACTTCAAGCTTCTTAAACATTC 3159
961 GTCAAGAGGTACAGCAATGATGGCTCAGCAACCACTTCAATACCTTACCAAGCAGC 1020
3160 GTCAAGAGGTACAGCAATGATGGCTCAGCAACCACTTCAATACCTTACCAAGCAGC 3219
1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTTCGGTACGTCTCTCGGCTCTGGCAGCCAG 1080
3220 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTTCGGTACGTCTCTCGGCTCTGGCAGCCAG 3279
1081 GGTGCTCTCTCGGCTTTCGGGAGAGTGTTCATGATTCGGCAATACGGCTACCTGAGC 1140
3280 GGTGCTCTCTCGGCTTTCGGGAGAGTGTTCATGATTCGGCAATACGGCTACCTGAGC 3339
1141 CTCAACAAATTCGAGCGGAGCGGTGGAGCTTCTATCTTTTACCTGGGATATTTCCCT 1200
3340 CTGAACAAACGGGAGTCAGGAGTGGAGCGCTTCTTCAATTTTACTGCTGGAGTACTTTCCT 3399
1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTTCAGCTTACACTTTTGGAGGAGTGCCT 1260

3400 TCTCAGATGCTGCTGCTACCGGAAACAACTTTTACCTTACGCTACACTTTTGGAGCGTCTCT 3459
1261 TTCCACAGCAGCTACCGGCACACAGCAGCGCTTGGACCGGCTGATGAATCTCTCATCGAC 1320
3460 TTCCACAGCAGCTACCGCTCAGACGAGCAGCTTGGACCGGCTCATGAATCTCTCATCGAC 3519
1321 CAATACCTGTATTACTGAAACAGAACTCAAAATCAGTCGGAGTGGCCCAAAACAAAGGAC 1380
3520 CAGTACCTGTATTACTGAGCAGAAACAAACACTCCAAAGTGGAAACCAACGAGTCAAGG 3579
1381 TTGCTGTTTACCGGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACCTGGTACCT 1440
3580 CTTCAAGTCTTCTCAGCGCCGAGCAGTACATTTCCGGACAGCTTAGGAACCTGGCTTCTCT 3639
1441 GGACCTCTGTTATTCGGCAGCAGCGCGTCTTCTAAAAACAAAAACAGAAACAAACAGCAAT 1500
3640 GGACCTCTGTTTACCGCAGCAGCGAGTATCAAGAGACATCTCGCGATACAAACAGTGA 3699
1501 TTTCCTGGACTGGTGTTCAAAAATATAACCTCAATGGCGGTGAATTCATCATCAACCT 1560
3700 TACTCGTGGACTGGAGCTACCAAGTACCACTCAATGGCAGAGACTCTCTGGTGAATCCG 3759
1561 GGCACCTGTATGGCTTCACAAAGCAGCAGAGCAAGTCTTTCCTCCCATGAGCGGTCTC 1620
3760 GGGCCCGCATGGCAAGCCACAAAGCAGATGAAGAAAGTCTTTCCTCAGAGCGGGT 3819
1621 ATGATTTTGGAAAGAGAGCGCCGAGCTTCAAAACACTTGCATTTGGCAATGTCTGATTT 1680
3820 CTCATCTTTGGAGCAGAGCTTCAGAGAAACAAATGTGAACATTTGAAAAGGTGATTT 3879
1681 ACAGCAGAGAGGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
3880 ACAGCAGAGAGGAAATCGGAACAACTTCCGTGGCTACGGAGCAGTATGTTCTGTGA 3939
1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCTTCGACCGGAGATGTGATGTCTATGGGA 1800
3940 TCTACCACTTCCAGAGAGGCAACAGACAGCTTACCGCAGATGTCAACACACAAGGC 3999
1801 GCATTTACCTGGCATGGTGGCAAGATAGAGACGTGTACTTCAGGCGTCCCAATTTGGGCC 1860
4000 GTTCTTCAGGATGGTCTGGCAGCAGAGATGTGATCTTCAGGGGCCCATCTGGGCA 4059
1861 AAAATTTCTCACAGAGATGGACATTTTCAACCGCTCTCTCTTATGGCGGCTTTGACTC 1920
4060 AAGATTTCCACACACGAGCAGCATTTTCAACCTCTCTCTCTCTCATGGGTGGATTCGACT 4119
1921 AAGAACCGCTCTCTCAGATCTCTCATCAAAACACGCTTCTCTCGGATCTCTCGGCG 1980
4120 AAACACCTCTCTCAGATCTCTCATCAAGAAACACCCCGGTACCTCGAATCTCTCGACC 4179
1981 GAGTTTTCAGCTACAAAGTTTGTCTTCAATCACCCTTATCTCCACAGGACA-AGTGAG 2039
4180 ACCTTCAGTGGCGGAAGTTTGTCTTCTTATCAGACAGTATCTCCACGAGCAGGCTCAG 4239
2040 TGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGA 2099
4240 CGTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAACCGTGAATCCCGAATTTCA 4299
2100 GTACACATCAATTTATGCAAAATCTGCCAAGTGTGATTTTACTGTGAGCAACATGACT 2159
4300 GTACACTTCCAACTACAAAGTCTGTTAATCGTGGACTTACCGTGGATCTAATGGCGT 4359
2160 TTATACTGAGCTCGCCCATTTGGCACCGCTTACCTTACCCGCTCCCTGT 2209
4360 GTATTTCAGAGCTTCGCCCATTTGGCACCAGATACCTGACTCGTGAATCTGT 4409

Search completed: November 29, 2005, 14:27:45

Job time : 320.186 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 08:46:14 ; Search time 1033.63 Seconds
(without alignments)
14243.261 Million cell updates/sec

Title: US-10-696-900-1_COPY_2223_4431
Perfect score: 2209
Sequence: 1 atggctgcgcagtggttatct.....ttaccttaccggtccctcgt 2209

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 4996997 seqs, 3332346308 residues 9993994
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*
14: Geneseqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2209	100.0	2211	3	AAD00777
2	2209	100.0	2211	14	Adz27052 Adeno-ass
3	2209	100.0	2211	14	Adz26929 Adeno-ass
4	2209	100.0	4347	13	Adw39398 Adeno-ass
5	2209	100.0	4347	14	Adz46594 HSV-AAV s
6	2209	100.0	4718	3	AAD00772 Adeno-ass
7	2209	100.0	4718	10	Adz76507 Adeno-ass
8	2209	100.0	4718	10	Adl13984 Adeno-ass
9	2209	100.0	4718	12	Adg39758 AAV-1 gen
10	2202.6	99.7	7447	8	Acc58477 Vector pF
11	2196.2	99.4	2211	14	Adz26928 Adeno-ass
12	2124	96.2	2214	14	Adz26930 Adeno-ass
13	2109.8	95.5	2211	14	Adz26931 Adeno-ass
14	2105	95.3	2211	14	Adz27053 Adeno-ass
15	2105	95.3	4239	13	Adw39402 Adeno-ass
16	2105	95.3	4239	14	Adz46598 HSV-AAV s
17	2105	95.3	4683	4	Aaf23749 AAV6 DNA
18	2105	95.3	4683	10	Adl13983 Adeno-ass
19	2105	95.3	4683	12	Adg39763 AAV-6 gen

20	2100.2	95.1	2211	14	Adz26932
21	1798	81.4	1800	3	AAD00778
22	1639	74.2	2214	14	Adz26890
23	1620	73.3	2214	14	Adz26891
24	1616.8	73.2	3122	10	Adz76546
25	1616.6	73.2	2214	14	Adz26887
26	1615.2	73.1	3117	10	Adz76544
27	1615.2	73.1	3122	10	Adz76543
28	1613.6	73.0	3121	10	Adz76545
29	1611.8	73.0	2214	14	Adz27063
30	1611.8	73.0	4721	10	Adz76502
31	1611.8	73.0	4721	12	Adg39764
32	1611.8	73.0	4721	14	Adv67509
33	1611.8	73.0	4721	14	Adz27030
34	1607	72.7	2214	14	Adz26885
35	1607	72.7	2217	14	Adz26869
36	1605.4	72.7	2214	14	Adz26886
37	1605.4	72.7	2217	14	Adz26892
38	1605.4	72.7	3127	10	Adz76516
39	1603.8	72.6	2214	14	Adz26882
40	1603.8	72.6	2217	14	Adz26875
41	1603	72.6	1605	3	AAD00779
42	1602.2	72.5	2214	14	Adz26889
43	1602.2	72.5	2217	14	Adz26865
44	1602.2	72.5	2217	14	Adz26866
45	1600.6	72.5	2217	14	Adz26867

ALIGNMENTS

RESULT 1

AAD00777
ID AAD00777 standard; DNA; 2211 BP.
XX
AC AAD00777;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 capsid protein VP1 DNA.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy; vaccine; transgene; VP1; ss.
XX
OS Adeno associated virus serotype 1.
XX
FH Key Location/Qualifiers
FT CDS 1..2211
FT /*tag= a
FT /product= "VP1 protein"
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US025694.
XX
PR 05-NOV-1998; 98US-0107114P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
PP WPI; 2000-376571/32.
XX
DR P-PSDB; AAY71167.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.
XX
PS Claim 10; Page 83-87; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA

CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antisense delivery vectors. The AAV-1 does not induce the
CC formation of neutralising antibodies specific to any serotype of AAV
CC hence is useful for transforming host cells, and in the preparation of a
CC medicament for the delivery of transgene to a host. The present sequence
CC is an AAV-1 DNA encoding a cap protein VP1 which is useful in the
CC production of recombinant viral vector for gene delivery
XX
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;

Query Match 100.0%; Score 2209; DB 3; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGCAATCGC 60
DB 1 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGCAATCGC 60

QY 61 GAGTGTGGGACTTGAAACCTCGAGCCCGAAGCCCAAGCCCAACAGCAAGAGCAGGAC 120
DB 61 GAGTGTGGGACTTGAAACCTCGAGCCCGAAGCCCGAAGCCCAACAGCAAGAGCAGGAC 120

QY 121 GACGGCCGGGGTCTGGTGTCTCTCGCTACAAAGTACCTCGGACCCCTTCAACGGACTCGAC 180
DB 121 GACGGCCGGGGTCTGGTGTCTCTCGCTACAAAGTACCTCGGACCCCTTCAACGGACTCGAC 180

QY 181 AAGGGGAGCCGCTCAACGGGGGAGCGAGCGGCCCTCGAGCAAGAGGCTACGAC 240
DB 181 AAGGGGAGCCGCTCAACGGGGGAGCGAGCGGCCCTCGAGCAAGAGGCTACGAC 240

QY 241 CAGCAGCTCAAGAGCGGGTGACATCCGTACCTCGGTATAAACACCGCCAGCCGAGTTT 300
DB 241 CAGCAGCTCAAGAGCGGGTGACATCCGTACCTCGGTATAAACACCGCCAGCCGAGTTT 300

QY 301 CAGGAGCGCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGGAGCAGCTCTTCCAG 360
DB 301 CAGGAGCGCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGGAGCAGCTCTTCCAG 360

QY 361 GCCAAGAAGCGGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGTCTCT 420
DB 361 GCCAAGAAGCGGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGTCTCT 420

QY 421 GGAAAGAAACGTCCGGTAGAGCAGTCGCCAAGAGCAGACCTCTCTCGGGCATCGGC 480
DB 421 GGAAAGAAACGTCCGGTAGAGCAGTCGCCAAGAGCAGACCTCTCTCGGGCATCGGC 480

QY 481 AAGACAGGCGCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTCGGCACTCAGAG 540
DB 481 AAGACAGGCGCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTCGGCACTCAGAG 540

QY 541 TCAGTCCCCGATCCCAAACTCTCGGAGAACCTTCAGAGCAACCCCGCTGTGTGGACCT 600
DB 541 TCAGTCCCCGATCCCAAACTCTCGGAGAACCTTCAGAGCAACCCCGCTGTGTGGACCT 600

QY 601 ACTTACAATGGCTTCAGGCGGTGGCGCAGCAATGGCAGACAATAACGAAGCGCCGACGGA 660
DB 601 ACTTACAATGGCTTCAGGCGGTGGCGCAGCAATGGCAGACAATAACGAAGCGCCGACGGA 660

QY 661 GTGGGTAAATGCCCTCAGGAAATTTGGCATTTGCGATTCCACATGGCTGGCGGACAGAGTCATC 720
DB 661 GTGGGTAAATGCCCTCAGGAAATTTGGCATTTGCGATTCCACATGGCTGGCGGACAGAGTCATC 720

QY 721 ACCACAGCAACCCGCACTCTGGGCTTGGCCCACTCAATAACCACTCTACAGAAATC 780
DB 721 ACCACAGCAACCCGCACTCTGGGCTTGGCCCACTCAATAACCACTCTACAGAAATC 780

QY 781 TCCAGTGTCTTCAACGGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTCGG 840

DB 781 TCCAGTGTCTCAACGGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTCGG 840

QY 841 GGGTATTTTGTATTTCAACAGATTCACATGCGCACTTTTCCACAGTGATGGCAGGACTC 900

DB 841 GGGTATTTTGTATTTCAACAGATTCACATGCGCACTTTTCCACAGTGATGGCAGGACTC 900

QY 901 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960

DB 901 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960

QY 961 GTCAAGGAGGTCACCAACGAATGATGGCGTCACAAACCATCGCTAATAAATTTACCAAGCAG 1020

DB 961 GTCAAGGAGGTCACCAACGAATGATGGCGTCACAAACCATCGCTAATAAATTTACCAAGCAG 1020

QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCGGTTCGGTCTCGGCTCTCGGACCCAG 1080

DB 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCGGTTCGGTCTCGGCTCTCGGACCCAG 1080

QY 1081 GGCTGCCCTCCCTCGGTTCCGGCGGACGTGTTTCAATGATTCGCAATACGGCTACCTGACG 1140

DB 1081 GGCTGCCCTCCCTCGGTTCCGGCGGACGTGTTTCAATGATTCGCAATACGGCTACCTGACG 1140

QY 1141 CTCAACAATGGCAGCAAGCCGTGGGAGCTTCACTCTTTTACTGCTTGGAAATTTTCCCT 1200

DB 1141 CTCAACAATGGCAGCAAGCCGTGGGAGCTTCACTCTTTTACTGCTTGGAAATTTTCCCT 1200

QY 1201 TCTCAGATGCTGAGAAACGGGCAACATTTTACCTTCAGCTACACCTTTGAGGAAGTGCT 1260

DB 1201 TCTCAGATGCTGAGAAACGGGCAACATTTTACCTTCAGCTACACCTTTGAGGAAGTGCT 1260

QY 1261 TTCCAAGAGCTACCGGCAAGCCGAGCTGGAGCCGGCTGATGAATCTCTCATCGAC 1320

DB 1261 TTCCAAGAGCTACCGGCAAGCCGAGCTGGAGCCGGCTGATGAATCTCTCATCGAC 1320

QY 1321 CAATACCTGTATTACTCTGAAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAGGAC 1380

DB 1321 CAATACCTGTATTACTCTGAAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAGGAC 1380

QY 1381 TTGCTGTTTGTAGCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACCTGGTACCT 1440

DB 1381 TTGCTGTTTGTAGCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACCTGGTACCT 1440

QY 1441 GGACCTGTATTCCGAGCAGCGCGTTTCTAAACAAAGAGAGCAAGCAAGAGCAGCAAT 1500

DB 1441 GGACCTGTATTCCGAGCAGCGCGTTTCTAAACAAAGAGAGCAAGCAAGAGCAGCAAT 1500

QY 1501 TTTTACCTGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATTCATCATCAACCT 1560

DB 1501 TTTTACCTGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATTCATCATCAACCT 1560

QY 1561 GGCACTGTATGGCTCTACAAAGAGAGCAAGAGTCTTCTTCCCATGAGCGGTGTC 1620

DB 1561 GGCACTGTATGGCTCTACAAAGAGAGCAAGAGTCTTCTTCCCATGAGCGGTGTC 1620

QY 1621 ATGATTTTGGAAAGAGAGCGCCCGAGCTTCAAAACATGCAATTTGGACATGTCATGATT 1680

DB 1621 ATGATTTTGGAAAGAGAGCGCCCGAGCTTCAAAACATGCAATTTGGACATGTCATGATT 1680

QY 1681 ACAGCAGAGAGAGAAATTAAGCCACTAACCTCTGAGCGGCAAGAGATTTGGGACCGTG 1740

DB 1681 ACAGCAGAGAGAGAAATTAAGCCACTAACCTCTGAGCGGCAAGAGATTTGGGACCGTG 1740

QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCTCGCAGCGGAGATGTGATGCTATGGGA 1800

DB 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCTCGCAGCGGAGATGTGATGCTATGGGA 1800

QY 1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTGGGCG 1860

DB 1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTGGGCG 1860

QY 1861 AAAATTTCTCACAGATGGACACTTTTCAACCCGTCTCTCTTATGGGCGGCTTTGGACATC 1920

Db 1861 AAAATTCTCACACAGATGGACATTTTCAACCGTCTCTCTTTATGGCGGCTTTGGACTC 1920
QY 1921 AAGAACCCGCTCTCTCAGATCTCTATCAAAAACACGCGCTGTCTCGAATCTCCGGCG 1980
Db 1921 AAGAACCCGCTCTCTCAGATCTCTATCAAAAACACGCGCTGTCTCGAATCTCCGGCG 1980
QY 1981 GAGTTTTCAGTACAAAGTTTGTCTCATTCACCCCAATATCTCCACAGCAAGTGAGT 2040
Db 1981 GAGTTTTCAGTACAAAGTTTGTCTCATTCACCCCAATATCTCCACAGCAAGTGAGT 2040
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAGTGCAG 2100
Db 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAGTGCAG 2100
QY 2101 TACACATCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGTGACAAACATGGACTT 2160
Db 2101 TACACATCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGTGACAAACATGGACTT 2160
QY 2161 TATACTGAGCCTGCGCCCAATTTGGCACCCGTTTACCTTACCGTCCCTGT 2209
Db 2161 TATACTGAGCCTGCGCCCAATTTGGCACCCGTTTACCTTACCGTCCCTGT 2209
RESULT 2
AD227052
ID AD227052 standard; DNA; 2211 BP.
XX
AC AD227052;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus DNA SEQ ID NO 202.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
PN W02005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004MO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX
PR (UYPE-) UNIV PENNSYLVANIA.
PA
XX
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
PI WPI; 2005-285437/29.
XX
DR
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
PT
XX
XX Disclosure; SEQ ID NO 202; 569pp; English.
PS
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative

CC colitis. The present sequence represents an adeno-associated virus DNA.
XX
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;
Query Match 100.0%; Score 2209; DB 14; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC 60
Db 1 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC 60
QY 61 GAGTGTGGGACTTGAACCTGGAGCCCGAAGCCCAAGCCCAACAGCAAGAGCAGGAC 120
Db 61 GAGTGTGGGACTTGAACCTGGAGCCCGAAGCCCAAGCCCAACAGCAAGAGCAGGAC 120
QY 121 GACGCCGGGGTCTGGTCTCTCGCTACAAGTACCTCGGACCCCTTCAACGGACTTCGAC 180
Db 121 GACGCCGGGGTCTGGTCTCTCGCTACAAGTACCTCGGACCCCTTCAACGGACTTCGAC 180
QY 181 AAGGGGAGCCCGTCAACGCGCGGAGCGAGCGGCGCTCGAGCAGCAAGGCTTACGAC 240
Db 181 AAGGGGAGCCCGTCAACGCGCGGAGCGAGCGGCGCTCGAGCAGCAAGGCTTACGAC 240
QY 241 CAGCAGCTCAAAAGCGGGTGAATCCGTTACCTTCGGGTATAACACGCCGCGGAGTTT 300
Db 241 CAGCAGCTCAAAAGCGGGTGAATCCGTTACCTTCGGGTATAACACGCCGCGGAGTTT 300
QY 301 CAGGAGCGTCTCAAGAGATACGCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG 360
Db 301 CAGGAGCGTCTCAAGAGATACGCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG 360
QY 361 GCCAAGAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGTCTCT 420
Db 361 GCCAAGAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGTCTCT 420
QY 421 GAAAGAAAAGCTCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTGGGCATCGGC 480
Db 421 GAAAGAAAAGCTCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGCATCGGC 480
QY 481 AAGACAGCCAGCAGCCGCTAAAGAGACTCAATTTTGGTTCAGACTGCGGACTCAGAG 540
Db 481 AAGACAGCCAGCAGCCGCTAAAGAGACTCAATTTTGGTTCAGACTGCGGACTCAGAG 540
QY 541 TCAGTCCCCGATCCACAACCTCTCGGAGAACCTTCCAGCAACCCCGCTCTGTGGGACCT 600
Db 541 TCAGTCCCCGATCCACAACCTCTCGGAGAACCTTCCAGCAACCCCGCTCTGTGGGACCT 600
QY 601 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAGCGCGCAGGA 660
Db 601 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAGCGCGCAGGA 660
QY 661 GTGGGTAATGCTTCAGGAAATTTGGCAATTCGATTTCCATGCTGGGCGCAGAGTATC 720
Db 661 GTGGGTAATGCTTCAGGAAATTTGGCAATTCGATTTCCATGCTGGGCGCAGAGTATC 720
QY 721 ACCACAGCAGCCGACCTTGGCCCTTGGCCACCTCAATAAACCCCTCTTACAGCAATC 780
Db 721 ACCACAGCAGCCGACCTTGGCCCTTGGCCACCTCAATAAACCCCTCTTACAGCAATC 780
QY 781 TCCAGTGTCTTCAACGGGGCCAGCAACGACCACTACTTGGGTACAGCAGCCCTGG 840
Db 781 TCCAGTGTCTTCAACGGGGCCAGCAACGACCACTACTTGGGTACAGCAGCCCTGG 840
QY 841 GGGTATTTGATTTCAACAGATTCCACTGCCATTTTCCAGTCTGGCAGCCACTC 900
Db 841 GGGTATTTGATTTCAACAGATTCCACTGCCATTTTCCAGTCTGGCAGCCACTC 900
QY 901 ATCAACAACTTTGGGGATTTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960
Db 901 ATCAACAACTTTGGGGATTTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960
QY 961 GTCAAGGAGGTCAACGAGATGATGGCGTCAACCACTCGCTAATAACCTTACAGCAGC 1020

Db 961 GTACAGGAGGTACGACGAATGATGGCGTCACAAACCATCGCTAATAACCTTACCGACG 1020
QY 1021 GTTCAAGTCTTCGAGCTCGAGATACAGCTTCGCTGCTCGCTCGCTCGGACACGAG 1080
Db 1021 GTTCAAGTCTTCGAGCTCGAGATACAGCTTCGCTGCTCGCTCGCTCGGACACGAG 1080
QY 1081 GGTGCTCCCTCCGTTCCCGGGGAGCGTTCATGATTTCCGCAATACGGCTACCTGAGC 1140
Db 1081 GGTGCTCCCTCCGTTCCCGGGGAGCGTTCATGATTTCCGCAATACGGCTACCTGAGC 1140
QY 1141 CTCAACAATGGCAGCCGAGCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTCCCT 1200
Db 1141 CTCAACAATGGCAGCCGAGCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTCCCT 1200
QY 1201 TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
Db 1201 TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
QY 1261 TTCCACAGCAGCTACGCGCACAGCCAGCGCTGGACCGGCTGATGATCTCTCATCGAC 1320
Db 1261 TTCCACAGCAGCTACGCGCACAGCCAGCGCTGGACCGGCTGATGATCTCTCATCGAC 1320
QY 1321 CAATACCTGTTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGAC 1380
Db 1321 CAATACCTGTTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGAC 1380
QY 1381 TTGCTGTTTACGCTGGGCTTCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACCT 1440
Db 1381 TTGCTGTTTACGCTGGGCTTCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACCT 1440
QY 1441 GGACCTGTTTACGCGCAGCGCGTCTTCTTAAACAAACAAACAGACAAACAGCAAT 1500
Db 1441 GGACCTGTTTACGCGCAGCGCGTCTTCTTAAACAAACAAACAGACAAACAGCAAT 1500
QY 1501 TTTTACCTGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAACCT 1560
Db 1501 TTTTACCTGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAACCT 1560
QY 1561 GGCACCTGCTATGCGCTTCACAAAGACGACGAAGTCTTTTCCCATGAGCGGTGTC 1620
Db 1561 GGCACCTGCTATGCGCTTCACAAAGACGACGAAGTCTTTTCCCATGAGCGGTGTC 1620
QY 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCATTTGGACAAATGCTATGATT 1680
Db 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCATTTGGACAAATGCTATGATT 1680
QY 1681 ACAGACGAAGAGAAATTAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGACCGTG 1740
Db 1681 ACAGACGAAGAGAAATTAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGACCGTG 1740
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTCATGCTATGGGA 1800
Db 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTCATGCTATGGGA 1800
QY 1801 GCATTAACCTGGCATGCTGGCAAGATAGAGCTGTAAGCTGCTGAGGCTCCATTTGGGCG 1860
Db 1801 GCATTAACCTGGCATGCTGGCAAGATAGAGCTGTAAGCTGCTGAGGCTCCATTTGGGCG 1860
QY 1861 AAAATTCCTCACACAGATGGACACTTTTACCCGCTCTCTTATGGGCGGCTTTGACATC 1920
Db 1861 AAAATTCCTCACACAGATGGACACTTTTACCCGCTCTCTTATGGGCGGCTTTGACATC 1920
QY 1921 AGAACCCTGCTCTCAGATCTCATCAAAAACACGCGCTGTTCTCGGAATCTCCGGCG 1980
Db 1921 AGAACCCTGCTCTCAGATCTCATCAAAAACACGCGCTGTTCTCGGAATCTCCGGCG 1980
QY 1981 GAGTTTTCAGCTACAAGTTTGGCTTCAATCATCAACCAATCTCCACAGCAAGTGAGT 2040
Db 1981 GAGTTTTCAGCTACAAGTTTGGCTTCAATCATCAACCAATCTCCACAGCAAGTGAGT 2040
QY 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGGAATCCCGAAGTGCAG 2100

Db 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGGAATCCCGAAGTGCAG 2100
QY 2101 TACACATCAATTTATGAAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACATGACTT 2160
Db 2101 TACACATCAATTTATGAAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACATGACTT 2160
QY 2161 TATACGTAGCCTCGGCCCATTTGGCACCCGTTTACCTTTACCCGTCCTCTGT 2209
Db 2161 TATACGTAGCCTCGGCCCATTTGGCACCCGTTTACCTTTACCCGTCCTCTGT 2209
RESULT 3
ADZ26929
ID ADZ26929 standard; DNA; 2211 BP.
XX
AC ADZ26929;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus DNA SEQ ID NO 79.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antitumor; dermatological; ds.
XX
OS Adeno-associated virus.
XX
PN WO2005033321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004WO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
PR 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX
WPI; 2005-285437/29.
XX
PT New adeno-associated virus (AAV) clade comprising at least three AAV
members, useful for preventing and/or treating arthritis, multiple
sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
infection and cancer.
XX
PS Disclosure; SEQ ID NO 79; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
at least three AAV members, where each member of the AAV clade is
phylogenetically related as determined using a Neighbor-Joining heuristic
by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
correction distance measurement of no more than 0.05. The methods and
compositions of the present invention are useful for the prevention
and/or treatment of rheumatoid arthritis, multiple sclerosis,
sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
colitis. The present sequence represents an adeno-associated virus DNA.
XX
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;
Query Match 100.0%; Score 2209; DB 14; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTGCCGATGGTATCTTCCAGATTTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60
Db 1 ATGGCTGCCGATGGTATCTTCCAGATTTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60

Db 3461 GGACCCCTGTTATTCGGCAGCAGCGCGTTTCTAATAACAAAAACAGACAACAAACAGCAAT 3520
QY 1501 TTTACTCGACTGGTCTTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCT 1560
Db 3521 TTTACTCGACTGGTCTTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCT 3580
QY 1561 GGCATGCTATGCGCTTCACAAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGTC 1620
Db 3581 GGCATGCTATGCGCTTCACAAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGTC 3640
QY 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCACACATGCAATGGACAATGTCATGATT 1680
Db 3641 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCACACATGCAATGGACAATGTCATGATT 3700
QY 1681 ACAGACGAAGAGAAATTAAGCCACTTAACCCCTGTGGCCACCAAGAAAGATTGGGACCGTG 1740
Db 3701 ACAGACGAAGAGAAATTAAGCCACTTAACCCCTGTGGCCACCAAGAAAGATTGGGACCGTG 3760
QY 1741 CGAGTCAATTTCCAGAGCAGCAGCAGACAGACCTTCGACCGGAGATGTGCATGTATGGGA 1800
Db 3761 CGAGTCAATTTCCAGAGCAGCAGCAGACAGACCTTCGACCGGAGATGTGCATGTATGGGA 3820
QY 1801 GCATTACCTGGCATGTGTGGCAAGATAGAGAGTGTACCTGCAGGGTCCCATTTGGGCC 1860
Db 3821 GCATTACCTGGCATGTGTGGCAAGATAGAGAGTGTACCTGCAGGGTCCCATTTGGGCC 3880
QY 1861 AAAATTCCTCACACAGATGGACACTTTACCCGCTCTCTTATGGCGGCTTTGGACTC 1920
Db 3881 AAAATTCCTCACACAGATGGACACTTTACCCGCTCTCTTATGGCGGCTTTGGACTC 3940
QY 1921 AAGAACCCTGCTCCTCAGATCCTCATCAAAAACACGCTGTTCCTCGAATCCTCCGGCG 1980
Db 3941 AAGAACCCTGCTCCTCAGATCCTCATCAAAAACACGCTGTTCCTCGAATCCTCCGGCG 4000
QY 1981 GAGTTTTCAGTACAAAGTTTGCTTCATTCATCACCCTATCTCCACAGACAAAGTGA 2040
Db 4001 GAGTTTTCAGTACAAAGTTTGCTTCATTCATCACCCTATCTCCACAGACAAAGTGA 4060
QY 2041 GTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100
Db 4061 GTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 4120
QY 2101 TACACATCAATTTATGCAAAATTCGCCAAGCTTGAATTTTACTGTGGACAACATGACTT 2160
Db 4121 TACACATCAATTTATGCAAAATTCGCCAAGCTTGAATTTTACTGTGGACAACATGACTT 4180
QY 2161 TATACTGAGCTCGCCCAATTTGGACCGGTTACCTTACCGTCCCTGT 2209
Db 4181 TATACTGAGCTCGCCCAATTTGGACCGGTTACCTTACCGTCCCTGT 4229

RESULT 5
ID AD246594
XX AD246594 standard; DNA; 4347 BP.

AC AD246594;

XX 30-JUN-2005 (first entry)

DT HSV-AAV sequence rep2cap1 for AAV/HSV vectors.

DE Genetic engineering; gene therapy; vector; cancer; neoplasm; cytostatic;
XX respiratory disease; respiratory-gen.; neurological disease;
KW neuroprotective; db.

XX Adeno-associated virus.
OS Human herpesvirus 1.

XX WO2005035743-A1.

XX 21-APR-2005.

XX 15-OCT-2003; 2003WO-CN000861.

XX 15-OCT-2003; 2003WO-CN000861.
PR (AGTC-) AGTC GENE TECHNOLOGY CO LTD.
XX Wu X, Cao H, Dong X;
PI WPI; 2005-296279/30.
XX Large-scale production, isolation and purification of serotype adeno-
PT associated virus vectors by infecting cells with recombinant herpes
PT simplex virus vectors, for use in gene therapy of e.g. cancer and
PT respiratory diseases.
XX Claim 1; SEQ ID NO 1; 79pp; Chinese.
XX The invention relates to a recombinant herpes simplex virus (HSV)
CC characterized in that its genome is inserted with a DNA sequence selected
CC from nucleotide sequences of AD246594-AD246598 or their homologous
CC sequences. Also included are preparing the recombinant HSV (comprising
CC the construction of DNA fragments that contains sequences AD246594-
CC AD246598, and respectively inserting these 5 DNA fragments into genome of
CC HSV by applying genetic engineering to give the recombinant HSV), large-
CC scale production of 5 serotype recombinant adeno-associated viruses (AAV)
CC 1, 3, 4, 5, and 6 (by preparing the recombinant HSV, establishing a one-
CC strain vector cell i.e. recombinant AAV vector cell-line, infecting the
CC recombinant AAV under the infective effect of the 5 serotype recombinant
CC the vector cell-line, separating and purifying the 5 serotype recombinant
CC AAV after lysing the AAV-containing cells and culture liquor to give a
CC crude lysate and further purification of rAAV by density-gradient
CC centrifugation or affinity chromatography) and a recombinant vector
CC plasmid pSNV-NX (containing ITR at both ends of AAV-1, AAV-3, AAV-4, AAV
CC -5 or AAV-6 genome, with immediate-early enhancer and promoter of
CC cytomegalovirus, and a polyclonal site and a polyA signal, respectively,
CC between the ITR (inverted terminal repeat), and neomycin-resistance gene-
CC expressing cassette at outer edge of ITR). The DNA sequence is inserted
CC into XbaI site in UL2 or UL44 gene in HSV genome. The DNA sequence of
CC sequence AD246594-AD246598 can also be inserted into other non-essential
CC gene regions in HSV genome. The recombinant HSV is optionally inserted
CC with other DNA sequences homologous with the already-specified fragments.
CC The vectors are for use in gene therapy of e.g. cancer, respiratory
CC diseases and neural diseases. The virus vectors are safe, with long
CC expression time and wide-spectrum of cell infection, even non-cleaved
CC cells and reverse axonal conduction through the incorporated HSV vectors,
CC and high transfer efficiently. The present sequence is the HSV-AAV
CC sequence rep2cap1 for the AAV/HSV vectors of the invention.

XX Sequence 4347 BP; 1106 A; 1186 C; 1142 G; 913 T; 0 U; 0 Other;

Query Match 100.0%; Score 2209; DB 14; Length 4347;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTGCG 60
Db 2021 ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTGCG 2080

QY 61 GAGTGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCCCAACAGCAAGAGCAGGAC 120
Db 2081 GAGTGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCCCAACAGCAAGAGCAGGAC 2140
QY 121 GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAAGTACCTCGGACCTTCAACGACTCGAC 180
Db 2141 GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAAGTACCTCGGACCTTCAACGACTCGAC 2200
QY 181 AAGGGGGAGCCCGTCAACGCGCGGACGCGAGCGGCCCTCGAGCAACAGGCTTACGAC 240
Db 2201 AAGGGGGAGCCCGTCAACGCGCGGACGCGAGCGGCCCTCGAGCAACAGGCTTACGAC 2260

QY 241 CAGCAGCTCAAGCGGGTGACAATCCGTACCTCGGTATTAACACCCGCGAGTTT 300
Db 2261 CAGCAGCTCAAGCGGGTGACAATCCGTACCTCGGTATTAACACCCGCGAGTTT 2320

QY 301 CAGGAGCGTCTCAAGAGATAGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
DB 2321 CAGGAGCGTCTCAAGAGATAGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2380
QY 361 GCCAAGACGGGTCTCGAAGCTCTCGGTCTGGTTGAGGAGGGCGCTAAGACGGCTCCT 420
DB 2381 GCCAAGACGGGTCTCGAAGCTCTCGGTCTGGTTGAGGAGGGCGCTAAGACGGCTCCT 2440
QY 421 GGAAAGAAACGTCCGGTAGAGAGTCCGCCAAGAGCAGACTCTCTCTCGGGCAGTCGGC 480
DB 2441 GGAAAGAAACGTCCGGTAGAGAGTCCGCCAAGAGCAGACTCTCTCTCGGGCAGTCGGC 2500
QY 481 AAGACAGGCGACGAGCCCTCAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540
DB 2501 AAGACAGGCGACGAGCCCTCAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 2560
QY 541 TCAGTCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTCTGTGGGACCT 600
DB 2561 TCAGTCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTCTGTGGGACCT 2620
QY 601 ACTAATAGCTTCAGCGGTGGCGCAACCAATGGCAGACAATAAGAGGGCGCGACGGA 660
DB 2621 ACTAATAGCTTCAGCGGTGGCGCAACCAATGGCAGACAATAAGAGGGCGCGACGGA 2680
QY 661 GTGGGTAAATGCCTCAGGAATTTGGCATTTGGCATTTCCAGATGGCTGGCGCAGAGTCATC 720
DB 2681 GTGGGTAAATGCCTCAGGAATTTGGCATTTGGCATTTCCAGATGGCTGGCGCAGAGTCATC 2740
QY 721 ACCACGACACCGCACCTGGGCTTTGCCACCTCAATAAACCACCTCTACAAGCAATC 780
DB 2741 ACCACGACACCGCACCTGGGCTTTGCCACCTCAATAAACCACCTCTACAAGCAATC 2800
QY 781 TCAGTGTCTTAACGGGGCGAGCAACGAACCACTATTGGGTACAGCAACCCCGTGG 840
DB 2801 TCAGTGTCTTAACGGGGCGAGCAACGAACCACTATTGGGTACAGCAACCCCGTGG 2860
QY 841 GGGTATTTTGAATTTCAACAGATTCACCTGCCACTTTTACCAGTGACTGGCAGCGACTC 900
DB 2861 GGGTATTTTGAATTTCAACAGATTTCCACTGCCACTTTTACCAGTGACTGGCAGCGACTC 2920
QY 901 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960
DB 2921 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 2980
QY 961 GTCAAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTTAATAACTTTACAGCAG 1020
DB 2981 GTCAAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTTAATAACTTTACAGCAG 3040
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCGTAAGTCTCGGCTCTGGCAGCAG 1080
DB 3041 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCGTAAGTCTCGGCTCTGGCAGCAG 3100
QY 1081 GGCTGCTCTCCTCGGCTTCCGGCGGACGTGTTTCATGATTTCCGCAATACGGCTACCTGAGC 1140
DB 3101 GGCTGCTCTCCTCGGCTTCCGGCGGACGTGTTTCATGATTTCCGCAATACGGCTACCTGAGC 3160
QY 1141 CTCAACAAATGGCAGCAGCGTGGAGCGTTTCATCTTTTACTGCTGGAAATATTTCCCT 1200
DB 3161 CTCAACAAATGGCAGCAGCGTGGAGCGTTTCATCTTTTACTGCTGGAAATATTTCCCT 3220
QY 1201 TCTCAGATCTGAGAACGGGCAACATTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 1260
DB 3221 TCTCAGATCTGAGAACGGGCAACATTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 3280
QY 1261 TTCCACAGAGCTTACCGGCACAGCCAGAGCTGGACGGCTGATGATCTCTCATCGAC 1320
DB 3281 TTCCACAGAGCTTACCGGCACAGCCAGAGCTGGACGGCTGATGATCTCTCATCGAC 3340
QY 1321 CAATACCTGTATTTACTGAAACAGAACTCAAAATCAGTCCGGAGAGTCCCAAAACAAGGAC 1380
DB 3341 CAATACCTGTATTTACTGAAACAGAACTCAAAATCAGTCCGGAGAGTCCCAAAACAAGGAC 3400

QY 1381 TTGCTGTTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCCAGCCCAAAACCTGGCTACT 1440
DB 3401 TTGCTGTTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCCAGCCCAAAACCTGGCTACT 3460
QY 1441 GGACCTCTTATTCGGCAGCAGCGGTCTTAAACAAAACAGACAAACAAACAGCAAT 1500
DB 3461 GGACCTCTTATTCGGCAGCAGCGGTCTTAAACAAAACAGACAAACAAACAGCAAT 3520
QY 1501 TTTTACCTGGACTGTGGTCTTCAAAATATATAACCTCAATGGGGGTGAATCATCATCAACCT 1560
DB 3521 TTTTACCTGGACTGTGGTCTTCAAAATATATAACCTCAATGGGGGTGAATCATCATCAACCT 3580
QY 1561 GGCACGTCTATGGCTCTACACAAAGACGACGAAGCAAGTCTTCTTCCATGAGCGGTGC 1620
DB 3581 GGCACGTCTATGGCTCTACACAAAGACGACGAAGCAAGTCTTCTTCCATGAGCGGTGC 3640
QY 1621 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAACTGCTATTTGGCAATGTCTATGATT 1680
DB 3641 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAACTGCTATTTGGCAATGTCTATGATT 3700
QY 1681 ACAGACGAAGAGGAAATTAAGGCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
DB 3701 ACAGACGAAGAGGAAATTAAGGCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 3760
QY 1741 GCAGTCAATTTCCAGAGCAGCAGACAGACCTCGACCGGAGATGTGCATGTATGGGA 1800
DB 3761 GCAGTCAATTTCCAGAGCAGCAGCAGACCTCGACCGGAGATGTGCATGTATGGGA 3820
QY 1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCC 1860
DB 3821 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCC 3880
QY 1861 AAAATTCCTCACAGATGGACACATTTTCAACCGCTCTCTTTATGGGCGGCTTTGGACTC 1920
DB 3881 AAAATTCCTCACAGATGGACACATTTTCAACCGCTCTCTCTTTATGGGCGGCTTTGGACTC 3940
QY 1921 AAGAACCCGCTCTCAGATCTCATCAAAACACGCTGTTCCTCGCAATCCTCCGGCG 1980
DB 3941 AAGAACCCGCTCTCAGATCTCTCATCAAAACACGCTGTTCCTCGCAATCCTCCGGCG 4000
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTTCACTTCACTACCCCAATCTCCACAGGCAAGTGAGT 2040
DB 4001 GAGTTTTCAGCTACAAAGTTTGTCTTCACTTCACTACCCCAATCTCCACAGGCAAGTGAGT 4060
QY 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGAG 2100
DB 4061 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGAG 4120
QY 2101 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 2160
DB 4121 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 4180
QY 2161 TATACTGAGCTTCGCCCCCATTTGGCACCCGTTTACCTTTACCCGCTCCCTGT 2209
DB 4181 TATACTGAGCTTCGCCCCCATTTGGCACCCGTTTACCTTTACCCGCTCCCTGT 4229

RESULT 6

AAD00772

ID AAD00772 standard; DNA; 4718 BP.

XX AAD00772;

AC AAD00772;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 DNA.

DE Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;

KW cap protein; recombinant viral vector; gene delivery; gene therapy;

KW vaccine; transgene; ss.

XX Adeno associated virus serotype 1.

OS Adeno associated virus serotype 1.

```

FH Key Location/Qualifiers
FT repeat_unit 1..143
FT /tag= a
FT /label= 5' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT protein_bind 89..110
FT /tag= b
FT /bound moiety= "Rep protein"
FT misc_feature 124..125
FT /tag= c
FT /note= "terminal resolute site (TRS)"
FT misc_binding 219..226
FT /tag= d
FT /bound moiety= "USF"
FT promoter 236..299
FT /tag= e
FT /label= P5_promoter
FT protein_bind 237..245
FT /tag= f
FT /bound moiety= "YY1 factor"
FT TATA_signal 270..275
FT /tag= g
FT /label= P5_TATA-Box
FT misc_feature 299..306
FT /tag= h
FT /note= "YY1/p5 RNA"
FT CDS 335..2272
FT /tag= j
FT /product= "Rep 68"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "the coding region is interrupted by intron"
FT CDS 335..2206
FT /tag= i
FT /product= "Rep 78"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT TATA_signal 857..862
FT /tag= l
FT /label= P19_TATA_Box
FT misc_feature 882..883
FT /tag= m
FT /note= "P19 RNA"
FT CDS 1007..2272
FT /tag= o
FT /product= "Rep 40"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "the coding region is interrupted by intron"
FT CDS 1007..2206
FT /tag= n
FT /product= "Rep 52"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT TATA_signal 1836..1841
FT /tag= p
FT /label= P40_TATA-BOX
FT misc_feature 1875..1876
FT /tag= q
FT /note= "P40 RNA"
FT intron 1924..2220
FT /tag= k
FT /note= "This region interrupts the coding sequence of Rep
FT 68 and Rep 40"
FT CDS 2223..4433
FT /tag= r
FT /product= "vpi protein"
FT /note= "Capsid protein"
FT CDS 2634..4433
FT /tag= s
FT /product= "VP2 protein"

```

```

FT /note= "Capsid protein"
FT /partial
FT 2829..4433
FT /tag= t
FT /product= "vp3 protein"
FT /note= "Capsid protein"
FT 4447..4452
FT /tag= u
FT repeat_unit 4576..4718
FT /tag= v
FT /label= 3' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT WO200028061-A2.
FT 18-MAY-2000.
FT 02-NOV-1999; 99WO-US025694.
FT 05-NOV-1998; 98US-0107114P.
FT (UYPE-) UNIV PENNSYLVANIA.
FT Wilson JM, Xiao W;
FT WPI; 2000-376571/32.
FT P-PSDB; AAV71161, AAV71164, AAV71165, AAV71166, AAV71167, AAV71168,
FT AAV71169.
FT Novel adeno-associated virus serotype 1 polynucleotide useful for
FT preparation of medicament for delivery of a transgene to a host.
FT Claim 1; Fig 1; 108pp; English.
FT
FT The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
FT characterised by two inverted terminal repeats (ITR) and open reading
FT frames for rep and capsid (cap) proteins. The rep reading frame encodes
FT four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap encoding
FT frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
FT sequence or its fragments particularly ITRs, rep and cap coding regions,
FT are useful in production of recombinant viral vectors for gene delivery.
FT These vectors can be used as gene therapy vectors, vaccine vectors or
FT antisense delivery vectors. The AAV-1 does not induce the formation of
FT neutralising antibodies specific to any serotype of AAV hence is useful
FT for transforming host cells, and in the preparation of a medicament for
FT the delivery of transgene to a host
FT
FT SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
FT
FT Query Match 100.0%; Score 2209; DB 3; Length 4718;
FT Best Local Similarity 100.0%; Pred. No. 0;
FT Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
FT QY 1 ATGGCTGCCGANGGTTATCTTCAGATTGGCTCGAGACAACTCTCTGAGGCGATTCCG 60
FT |||||
FT 2223 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGACAACTCTCTGAGGCGATTCCG 2282
FT |||||
FT QY 61 GAGTGTGGGACTTGAACCTCGAGCCCCGAGCCCAAGCCCAACCCAGCAAGAGCAGGAC 120
FT |||||
FT 2283 GAGTGTGGGACTTGAACCTCGAGCCCCGAGCCCAAGCCCAACCCAGCAAGAGCAGGAC 2342
FT |||||
FT QY 121 GACGGCGGGGCTCTGGTCTTCTCGCTACAAGTACTCTCGAGCCCTTCAACGGACTCGAC 180
FT |||||
FT 2343 GACGGCGGGGCTCTGGTCTTCTCGCTACAAGTACTCTCGAGCCCTTCAACGGACTCGAC 2402
FT |||||
FT QY 181 AAGGGGAGCCCGCTCAACCGCGCGACCGAGCGGCCCTTCGAGCAGCAAGAGGCTACGAC 240
FT |||||
FT 2403 AAGGGGAGCCCGCTCAACCGCGCGACCGAGCGGCCCTTCGAGCAGCAAGAGGCTACGAC 2462
FT |||||
FT QY 241 CAGCAGCTCAAGCGGGGTGACATTCGTACTCTCGGTATATACCAACCGCGCGCGGCTTT 300
FT |||||
FT 2463 CAGCAGCTCAAGCGGGGTGACATTCGTACTCTCGGTATATACCAACCGCGCGCGGCTTT 2522
FT |||||

```

QY 301 CAGGAGCGTCTGCAAGAGATAGTCTCTTTTGGGGGCAACCTCGGGGCGAGCAGTCTTCCAG 360
Db 2523 CAGGAGCGTCTGCAAGAGATAGTCTCTTTTGGGGGCAACCTCGGGGCGAGCAGTCTTCCAG 2582
QY 361 GCCAAGAGCGGGTCTCTCAACCTCTCGGTCTGGTTGAGGAGGCGCTAAGACGGCTCCT 420
Db 2583 GCCAAGAGCGGGTCTCTCAACCTCTCGGTCTGGTTGAGGAGGCGCTAAGACGGCTCCT 2642
QY 421 GGAAAGAAACGTCCGGTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCGGCCATCGGC 480
Db 2643 GGAAAGAAACGTCCGGTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCGGCCATCGGC 2702
QY 481 AAGACAGGCGAGGCGCCCTTAAAGAGAGACTCAATTTTGGTTCAGACTGGGAGCTCAGAG 540
Db 2703 AAGACAGGCGAGGCGCCCTTAAAGAGAGACTCAATTTTGGTTCAGACTGGGAGCTCAGAG 2762
QY 541 TCAGTCCCGATCCACAACTCTCGGAGAACTCTCCAGCAACCCCGCTCTGTGGGACCT 600
Db 2763 TCAGTCCCGATCCACAACTCTCGGAGAACTCTCCAGCAACCCCGCTCTGTGGGACCT 2822
QY 601 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAAGAGGCGCCGACGGA 660
Db 2823 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAAGAGGCGCCGACGGA 2882
QY 661 GTGGGTAAATGCCCTCAGGAATAATGGCAATGGGATTCAGATGGCTGGGCGACAGATCATC 720
Db 2883 GTGGGTAAATGCCCTCAGGAATAATGGCAATGGGATTCAGATGGCTGGGCGACAGATCATC 2942
QY 721 ACCACAGCACCGGACCTTGGGCTTGGCCACCTCAATAACCACTCTTCAAGCAAAATC 780
Db 2943 ACCACAGCACCGGACCTTGGGCTTGGCCACCTCAATAACCACTCTTCAAGCAAAATC 3002
QY 781 TCAGTGTCTTCAACGGGGCGAGCAACGAAACCACTACTTGGGTACAGCACCCCTGG 840
Db 3003 TCAGTGTCTTCAACGGGGCGAGCAACGAAACCACTACTTGGGTACAGCACCCCTGG 3062
QY 841 GGGTATTTTGAATTTCAACAGATTCACCTGGCACTTTTACCAAGTGGGAGGAGTCTC 900
Db 3063 GGGTATTTTGAATTTCAACAGATTTCCACTGCCATTTTCCACAGTGGTGGGCGAGCTC 3122
QY 901 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
Db 3123 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 3182
QY 961 GTCAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTAATAACCTTACCAGCAG 1020
Db 3183 GTCAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTAATAACCTTACCAGCAG 3242
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACGACTTCCGTAGTCTCGGCTCTGGGACGAG 1080
Db 3243 GTTCAAGTCTTCTCGGACTCGGAGTACGACTTCCGTAGTCTCGGCTCTGGGACGAG 3302
QY 1081 GGTGCTCTCCCTCGGTTCCGGGGGAGCGTGTTCATGATTCGGCAATACGGCTACCTGAGC 1140
Db 3303 GGTGCTCTCCCTCGGTTCCGGGGGAGCGTGTTCATGATTCGGCAATACGGCTACCTGAGC 3362
QY 1141 CTCAACAAATGGCAGCAAGCGGTGGAGCGTTTCATCTTTTACTGCTGGAAATATTTCCCT 1200
Db 3363 CTCAACAAATGGCAGCAAGCGGTGGAGCGTTTCATCTTTTACTGCTGGAAATATTTCCCT 3422
QY 1201 TCTCAGATCTGAGAACGGGCAACACTTTTACCTTCAGCTACACTTTGAGGAGTGGCT 1260
Db 3423 TCTCAGATCTGAGAACGGGCAACACTTTTACCTTCAGCTACACTTTGAGGAGTGGCT 3482
QY 1261 TTCCACAGCAGCTACGCGACAGCCAGAGCTCGAGCGGTGATGATCTCTCTCATCGAC 1320
Db 3483 TTCCACAGCAGCTACGCGACAGCCAGAGCTCGAGCGGTGATGATCTCTCTCATCGAC 3542
QY 1321 CAATACCTGTATTTACTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGAC 1380
Db 3543 CAATACCTGTATTTACTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGAC 3602

QY 1381 TTGCTGTTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCCAGCCCAAAAATCTGGCTACT 1440
Db 3603 TTGCTGTTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCCAGCCCAAAAATCTGGCTACT 3662
QY 1441 GGACCTCTGTATTCGGCAGCAGCGCTTTCTAAAAAAGAGCAACAAACAGCAAT 1500
Db 3663 GGACCTCTGTATTCGGCAGCAGCGCTTTCTAAAAAAGAGCAACAAACAGCAAT 3722
QY 1501 TTTTACTGGACTGTGTCTTCAAAATATTAACCTCAATGGCGTGAATCATCATCAACCT 1560
Db 3723 TTTTACTGGACTGTGTCTTCAAAATATTAACCTCAATGGCGTGAATCATCATCAACCT 3782
QY 1561 GGCACCTGTATGGCCTCACAAAGACGAGCAAGCAAGTCTTTTCCCATGAGCGGTGTC 1620
Db 3783 GGCACCTGTATGGCCTCACAAAGACGAGCAAGCAAGTCTTTTCCCATGAGCGGTGTC 3842
QY 1621 ATGATTTTGGAAAAAGAGCGCCGAGCTTCAAAACACTGCAATTGGACAATGTCTATGATT 1680
Db 3843 ATGATTTTGGAAAAAGAGCGCCGAGCTTCAAAACACTGCAATTGGACAATGTCTATGATT 3902
QY 1681 ACAGCAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTTGGGACCGTG 1740
Db 3903 ACAGCAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTTGGGACCGTG 3962
QY 1741 GCGGTCAATTTCCAGCAGCAGCAGCAGACCTCGCAGCGGAGATGTGCATGTATGGGA 1800
Db 3963 GCGGTCAATTTCCAGCAGCAGCAGCAGACCTCGCAGCGGAGATGTGCATGTATGGGA 4022
QY 1801 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGACGGGTCCCATTTGGGCC 1860
Db 4023 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGACGGGTCCCATTTGGGCC 4082
QY 1861 AAAATTCCTCACAGATGGACATTTTCAACCGCTCTCTTTATGGCGGCTTTGGACTC 1920
Db 4083 AAAATTCCTCACAGATGGACATTTTCAACCGCTCTCTTTATGGCGGCTTTGGACTC 4142
QY 1921 AAGAACCCGCTCTCAGTCTCATCAAAACACGCTGTCTCTCGAATCCTCCGGCG 1980
Db 4143 AAGAACCCGCTCTCAGTCTCATCAAAACACGCTGTCTCTCGAATCCTCCGGCG 4202
QY 1981 GAGTTTTCAGCTTACAAAGTTTCTTCAATTCATCACCAATCTCCACAGGACAAGTGA 2040
Db 4203 GAGTTTTCAGCTTACAAAGTTTCTTCAATTCATCACCAATCTCCACAGGACAAGTGA 4262
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTCAG 2100
Db 4263 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTCAG 4322
QY 2101 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACATGGACTT 2160
Db 4323 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACATGGACTT 4382
QY 2161 TATACTGAGCCTCGCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGT 2209
Db 4383 TATACTGAGCCTCGCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGT 4431

RESULT 7

ADE76507

ID ADE76507 standard; DNA; 4718 BP.

XX ADE76507;

AC ADE76507;

XX 29-JAN-2004 (first entry)

XX Adeno-associated virus (AAV) related DNA, SEQ ID No 6.

DE adeno-associated virus; AAV; cytostatic; antipneumatic; antirheumatic;

XX antirheumatic; neuroprotective; antidiabetic; antithyroid;

KW dermatological; antiinflammatory; gene therapy; vaccine;

KW hyperproliferative; cancer; psoriasis; autoimmune disease;

KW rheumatoid arthritis; multiple sclerosis; diabetes;

KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; da.

XX Adeno-associated virus 1.
OS EP1310571-A2.
XX 14-MAY-2003.
XX 12-NOV-2002; 2002BP-00257826.
XX 13-NOV-2001; 2001US-0350607P.
PR 17-DEC-2001; 2001US-0341117P.
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386675P.
XX (UIPPE-) UNIV PENNSYLVANIA.
XX Gao G, Wilson JM, Alvira M;
XX WPI; 2003-450984/43.
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprising subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX Claim 14; SEQ ID NO 6; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antirheumatic, antiarthritic, neuroprotective,
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences may be used e.g. for preventing or treating hyperproliferative
CC conditions such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 2209; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTGCG 60
DB 2223 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTGCG 2282
QY 61 GAGTGTGGGACTTGAAACCTTGAGCCCGGAGCCCAAGCCCAACGCAAGGAGGAGGAC 120
DB 2283 GAGTGTGGGACTTGAAACCTTGAGCCCGGAGCCCAAGCCCAACGCAAGGAGGAGGAC 2342
QY 121 GAGCGCGGGTCTGGTCTTCTGCTCAAGTACTCGGACCTTCAACGCAATCGAC 180
DB 2343 GACGGCGGGGTCTGGTCTTCTGCTCAAGTACTCGGACCTTCAACGCAATCGAC 2402
QY 181 AAGGGGAGCCCGCTCAACCGCGGGAGCGAGCGGCCCTCGAGCAGCAAGGCTTACGAC 240
DB 2403 AAGGGGAGCCCGCTCAACCGCGGGAGCGAGCGGCCCTCGAGCAGCAAGGCTTACGAC 2462
QY 241 CAGCAGCTCAAGCGGGTGACATTCCTGATACCTCGGGTATTAACCGCCGACCGGAGTTT 300
DB 2463 CAGCAGCTCAAGCGGGTGACATTCCTGATACCTCGGGTATTAACCGCCGACCGGAGTTT 2522
QY 301 CAGGCGGTCTGCAGAGGATAGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360
DB 2523 CAGGCGGTCTGCAGAGGATAGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 2582
QY 361 GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTGGAGGAGCGCTAAGACGGCTCCT 420

DB 2583 GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGGAAGGCGCTAAGACGGCTCCT 2642
QY 421 GGAAAGAAACGTCGGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGCGCATCGCG 480
DB 2643 GGAAAGAAACGTCGGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGCGCATCGCG 2702
QY 481 AAGACAGGCCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG 540
DB 2703 AAGACAGGCCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG 2762
QY 541 TCAGTCCCGCATCCACACCTCTCGAGAACTCCAGCAACCCCGCTCTGTGGGACCT 600
DB 2763 TCAGTCCCGCATCCACACCTCTCGAGAACTCCAGCAACCCCGCTCTGTGGGACCT 2822
QY 601 ACTCAATGGCTTCAGGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGGA 660
DB 2823 ACTCAATGGCTTCAGGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGGA 2882
QY 661 GTGGGTAATGCTCAGGAAATTTGGCATTCGATTCACATGGCTGGGCGCAGAGTCATC 720
DB 2883 GTGGGTAATGCTCAGGAAATTTGGCATTCGATTCGATTCGCTTGGGCGCAGAGTCATC 2942
QY 721 ACCACGACCCGACCTGGGCGCTTGGCCACTACATAAACCACTCTTACAGCAAAATC 780
DB 2943 ACCACGACCCGACCTGGGCGCTTGGCCACTACATAAACCACTCTTACAGCAAAATC 3002
QY 781 TCCAGTGTCTTCAACGGGGCCAGCAACGACCACTACTTGGGCTTACAGCAACCCCTG 840
DB 3003 TCCAGTGTCTTCAACGGGGCCAGCAACGACCACTACTTGGGCTTACAGCAACCCCTG 3062
QY 841 GGGTATTTTGAATTTCAACAGATTTCCACTGCCACTTTTCCACGCTGAGTGGCAGGACTC 900
DB 3063 GGGTATTTTGAATTTCAACAGATTTCCACTGCCACTTTTCCACGCTGAGTGGCAGGACTC 3122
QY 901 ATCAACAACAAATTTGGGATTCGGGCCAAGAGACTCACTTCAAACTCTTCAACATCCAA 960
DB 3123 ATCAACAACAAATTTGGGATTCGGGCCAAGAGACTCACTTCAAACTCTTCAACATCCAA 3182
QY 961 GTCAAGGAGGTCAACGACGAATGATGGCGTCAACAACCATCGCTTAATAACCTTACGACG 1020
DB 3183 GTCAAGGAGGTCAACGACGAATGATGGCGTCAACAACCATCGCTTAATAACCTTACGACG 3242
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGATCTCGGCTCTGCGCAACAG 1080
DB 3243 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGATCTCGGCTCTGCGCAACAG 3302
QY 1081 GGTGCTCTCCTCGTTCGGGGGAGCGTTCATGATTTCCGCAATACGGCTACCTGACG 1140
DB 3303 GGTGCTCTCCTCGTTCGGGGGAGCGTTCATGATTTCCGCAATACGGCTACCTGACG 3362
QY 1141 CTCAACAATGGCAGCCCAAGCCGTTGGGACGTTTCACTCTTTTACTGCTGGAATATTTCCCT 1200
DB 3363 CTCAACAATGGCAGCCCAAGCCGTTGGGACGTTTCACTCTTTTACTGCTGGAATATTTCCCT 3422
QY 1201 TCTCAGATCTCGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 1260
DB 3423 TCTCAGATCTCGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 3482
QY 1261 TTCCACAGCAGTTCACGCGCACAGCCGAGCTGGACCGGCTGATGTAATCTCTCATCGAC 1320
DB 3483 TTCCACAGCAGTTCACGCGCACAGCCGAGCTGGACCGGCTGATGTAATCTCTCATCGAC 3542
QY 1321 CAATACCTGATTTACTCTGAACAGACTCAAAATCAGTCCGGAGTCCCAAAACAAGGAC 1380
DB 3543 CAATACCTGATTTACTCTGAACAGACTCAAAATCAGTCCGGAGTCCCAAAACAAGGAC 3602
QY 1381 TTGCTGTTAGCGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAATCTGGCTACCT 1440
DB 3603 TTGCTGTTAGCGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAATCTGGCTACCT 3662
QY 1441 GGAACCTGTTATCGGCGAGCGGCTTTCTAAACAAAAACAGACAAACAAACGAAT 1500

Db 3663 GGACCCTGTTATCGGCAGCAGCGGCTTTCTAAAAAAGAGACAAACAGACAGCAAT 3722
QY 1501 TTTACTGTGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTT 1560
Db 3723 TTTACTGTGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTT 3782
QY 1561 GGCACCTGCTATGGCTTCACAAAAGACGAGCAAGTCTTTCCCATGAGCGGTGC 1620
Db 3783 GGCACCTGCTATGGCTTCACAAAAGACGAGCAAGTCTTTCCCATGAGCGGTGC 3842
QY 1621 ATGATTTTTGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATTTGGACAATGTCATGATT 1680
Db 3843 ATGATTTTTGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATTTGGACAATGTCATGATT 3902
QY 1681 ACAGACGAAGAGAAATTAAGCCACTAACCCTGTGGCCACCGAAGATTTGGACCGGTG 1740
Db 3903 ACAGACGAAGAGAAATTAAGCCACTAACCCTGTGGCCACCGAAGATTTGGACCGGTG 3962
QY 1741 GCAGTCAATTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGCATGTGGGA 1800
Db 3963 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCATGTGGGA 4022
QY 1801 GCATTAACCTGGCATGTGGCAAGATAGAGAGCTGTACTGCAGGGTCCCAATTTGGGCC 1860
Db 4023 GCATTAACCTGGCATGTGGCAAGATAGAGAGCTGTACTGCAGGGTCCCAATTTGGGCC 4082
QY 1861 AAAATTCCTCACACAGATGGACACTTTCAACCGGTCTCTCTTATGGCGGCTTTGACTC 1920
Db 4083 AAAATTCCTCACACAGATGGACACTTTCAACCGGTCTCTCTTATGGCGGCTTTGACTC 4142
QY 1921 AGAAGCCGCTCTCAGATCCTCATCAAAACAGCGCTGTCTCTGGAATCTCCGGCG 1980
Db 4143 AGAAGCCGCTCTCAGATCCTCATCAAAACAGCGCTGTCTCTGGAATCTCCGGCG 4202
QY 1981 GAGTTTTTCAGCTACAAAGTTTGTCTCATCATCAACCAATCTCCACAGACAAAGTGA 2040
Db 4203 GAGTTTTTCAGCTACAAAGTTTGTCTCATCATCAACCAATCTCCACAGACAAAGTGA 4262
QY 2041 GTGGAAATTAAGAGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAAGTGCAG 2100
Db 4263 GTGGAAATTAAGAGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAAGTGCAG 4322
QY 2101 TACACATCAATTAATCAAAATCTGCAAGCTTGTATTTACTGTGACACAAATGACTT 2160
Db 4323 TACACATCAATTAATCAAAATCTGCAAGCTTGTATTTACTGTGACACAAATGACTT 4382
QY 2161 TATACTGAGCTCGCCCCATTTGGCACCCGTTACCTTTACCCGTCCTCTGT 2209
Db 4383 TATACTGAGCTCGCCCCATTTGGCACCCGTTACCTTTACCCGTCCTCTGT 4431

RESULT 8

ADL13984

ID ADL13984 standard; DNA; 4718 BP.

XX

AC ADL13984;

XX

DT 06-MAY-2004 (first entry)

XX

DE Adeno-associated virus serotype 1 complete DNA.

XX

KW ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;

XX

KW expression construct; adeno-associated virus;

XX

KW integration efficiency element; inverted terminal repeat; integration;

XX

KW chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;

XX

KW retinoblastoma; inflammatory disease; arthritis;

XX

KW neurodegenerative disease.

XX

OS Adeno-associated virus 1.

XX

PN WO2003087334-A2.

XX

XX 23-OCT-2003.

XX

PF 09-APR-2003; 2003WO-US011191.

XX

PR 09-APR-2002; 2002US-0371044P.

XX

PA (CORR) CORNELL RES FOUND INC.

XX

PI Falck-Pedersen ES, Philipott N;

XX

DR WPI; 2003-833723/77.

XX

New expression construct comprising a nucleic acid sequence encoding an adeno-associated virus integration efficiency element, useful for treating cancer e.g. lung cancer or colon cancer or inflammatory disease e.g. arthritis.

XX

PS Disclosure; SEQ ID NO 3; 62pp; English.

XX

The invention relates to an expression construct comprising a nucleic acid sequence encoding an adeno-associated virus integration efficiency element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV ITRs) and site-specifically integrates into a host cell chromosome when provided to the host cell in conjunction with an AAV Rep protein. The expression construct can be used as a therapeutic factor for treating a mammal for a pathologic state which is cancer, including lung cancer, colon cancer, renal cancer, anal cancer, bile duct cancer, bladder cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer, cervical cancer, lymphoma, endometrial cancer, esophageal cancer, gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia, liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer, testicular cancer, thymus cancer or thyroid cancer. Other pathologic state includes inflammatory disease (arthritis), neurodegenerative disease, a disease of an organ attributed to the presence of increased or decreased level of a particular gene product(s). This sequence corresponds to the AAV serotype 1 complete DNA including the IEE sequence.

XX

SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;

QY

Query Match 100.0%; Score 2209; DB 10; Length 4718;

Db

Best Local Similarity 100.0%; Pred. No. 0;

QY

Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db

1 ATGGCTGGCATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60

QY

2223 ATGGCTGGCATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 2282

Db

61 GAGTGGTGGGACTTGAACCTGGAGCCCGGAAGCCCAAGCCAAAGCAAGCAGGAC 120

QY

2283 GAGTGGTGGGACTTGAACCTGGAGCCCGGAAGCCCAAGCCAAAGCAGGAC 2342

Db

121 GACGGCCGGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 180

QY

2343 GACGGCCGGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 2402

Db

181 AAGGGGGAGCCCGTCAACCGCGCGGACGCGCGGCTTCGACACGAAAGGCTACGAC 240

QY

2403 AAGGGGGAGCCCGTCAACCGCGCGGACGCGCGGCTTCGACACGAAAGGCTACGAC 2462

Db

241 CAGCAGCTCAAGCGGGTGACAATCCGTACCTCGGTATAACACCGCCGACCGGATTT 300

QY

2463 CAGCAGCTCAAGCGGGTGACAATCCGTACCTCGGTATAACACCGCCGACCGGATTT 2522

Db

301 CAGGAGCGTCTCAAGAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360

QY

2523 CAGGAGCGTCTCAAGAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 2582

Db

361 GCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGTCT 420

QY

2583 GCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGTCT 2642

Db

421 GGAAAGAAACGTCGCTAGAGCAGTCGCCCAAGAGCCAGACTCTCTCTCGGGCATCGGC 480

Db 2643 GGAAGAAACGTCGCTAGACAGTCGCGACAAAGACCCAGACTCCTCTCGGCACTCGGC 2702
QY |||||
Db 481 AAGACAGCCAGACCGCCGCTAAAGAGACCTCAATTTGGTTCAGACTGGCGACTCAGAG 540
QY |||||
Db 2703 AAGACAGCCAGACCGCCGCTAAAGAGACCTCAATTTGGTTCAGACTGGCGACTCAGAG 2762
QY |||||
Db 541 TCAGTCCCCGATCCCAACCTCTCGAGAACTCCAGCAACCCCGCTGCTGTGGACCT 600
QY |||||
Db 2763 TCAGTCCCCGATCCCAACCTCTCGAGAACTCCAGCAACCCCGCTGCTGTGGACCT 2822
QY |||||
QY 601 ACTACAAATGGCTTCAGAGCGGTGGCGCAACAATGGCAGACAATAACGAAGCGCCGACGGA 660
Db 2823 ACTACAAATGGCTTCAGAGCGGTGGCGCAACAATGGCAGACAATAACGAAGCGCCGACGGA 2882
QY |||||
QY 661 GTGGGTAAATGGCTTCAGAGAAATTTGGCAATTCAGATGCTGGCGAGACAGTCAATC 720
Db 2883 GTGGGTAAATGGCTTCAGAGAAATTTGGCAATTCAGATGCTGGCGAGACAGTCAATC 2942
QY |||||
QY 721 ACCACAGACCCGACCTGGGCTTTGCCACCTTACAATAACCACTCTTACAAGCAAAATC 780
Db 2943 ACCACAGACCCGACCTGGGCTTTGCCACCTTACAATAACCACTCTTACAAGCAAAATC 3002
QY |||||
QY 781 TCCAGTGTCTTCAACGGGGCCAGCAACGAACCACTACTTCGGCTTACAGCAACCCCTGG 840
Db 3003 TCCAGTGTCTTCAACGGGGCCAGCAACGAACCACTACTTCGGCTTACAGCAACCCCTGG 3062
QY |||||
QY 841 GGGTATTTGATTTCAACAGATTCACATGCTCCACTTTTACCAGCTGACTGGCAGCGACTC 900
Db 3063 GGGTATTTGATTTCAACAGATTCACATGCTCCACTTTTACCAGCTGACTGGCAGCGACTC 3122
QY |||||
QY 901 ATCAACAACAAATTTGGGATTCGCGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAA 960
Db 3123 ATCAACAACAAATTTGGGATTCGCGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAA 3182
QY |||||
QY 961 GTCAAGGAGGTCAACGAGATGATGGCGTCAACCAATCGCTTAATACCTTACAGCAGC 1020
Db 3183 GTCAAGGAGGTCAACGAGATGATGGCGTCAACCAATCGCTTAATACCTTACAGCAGC 3242
QY |||||
QY 1021 GTTCAAGTCTTTCGAGCTCGAGTACAGCTTCGATGCTCTGGCTCTGGCGACCAAG 1080
Db 3243 GTTCAAGTCTTTCGAGCTCGAGTACAGCTTCGATGCTCTGGCTCTGGCGACCAAG 3302
QY |||||
QY 1081 GGTGCTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGCTACCTGACG 1140
Db 3303 GGTGCTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGCTACCTGACG 3362
QY |||||
QY 1141 CTCAACAATGGCAGCAAGCCGTGGGACGTTTCACTCTTTTACTGCTGGAATATTTCCCT 1200
Db 3363 CTCAACAATGGCAGCAAGCCGTGGGACGTTTCACTCTTTTACTGCTGGAATATTTCCCT 3422
QY |||||
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 1260
Db 3423 TCTCAGATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 3482
QY |||||
QY 1261 TTCCACAGAGCTACCGCACAGCCAGAGCTCGACCGCTGATGATCTCTTCATCGAC 1320
Db 3483 TTCCACAGAGCTACCGCACAGCCAGAGCTCGACCGCTGATGATCTCTTCATCGAC 3542
QY |||||
QY 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC 1380
Db 3543 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC 3602
QY |||||
QY 1381 TTGCTGTTAGCGTGGGTCTCAGCTGGCATGTCTGTTGAGCCCAAAACTGGCTACCT 1440
Db 3603 TTGCTGTTAGCGTGGGTCTCAGCTGGCATGTCTGTTGAGCCCAAAACTGGCTACCT 3662
QY |||||
QY 1441 GGACCTGTATTCCGAGAGCGGCTTTCTTAAACAACCAACACACACACACACACAAAT 1500
Db 3663 GGACCTGTATTCCGAGAGCGGCTTTCTTAAACAACCAACACACACACACACAAAT 3722
QY |||||
QY 1501 TTTTACCTGGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCAACCCCT 1560
QY |||||

Db 3723 TTTTACCTGGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT 3782
QY |||||
QY 1561 GGCACCTGTATGGCTCTCACAAAAGACGACGAAGCAAAAGTTCTTTCCCATGAGCGGTGTC 1620
Db 3783 GGCACCTGTATGGCTCTCACAAAAGACGACGAAGCAAAAGTTCTTTCCCATGAGCGGTGTC 3842
QY |||||
QY 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACTGCACTGCAATTTGGACAATGTCAATTT 1680
Db 3843 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACTGCACTGCAATTTGGACAATGTCAATTT 3902
QY |||||
QY 1681 ACAGACGAGAGGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTTGGGACCCGTG 1740
Db 3903 ACAGACGAGAGGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTTGGGACCCGTG 3962
QY |||||
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACAGACCTCGCAGCGGAGATGTGCATCTATGGGA 1800
Db 3963 GCAGTCAATTTCCAGAGCAGCAGCAGACAGACCTCGCAGCGGAGATGTGCATCTATGGGA 4022
QY |||||
QY 1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACTTCAGGGTCCCATTTGGGCC 1860
Db 4023 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACTTCAGGGTCCCATTTGGGCC 4082
QY |||||
QY 1861 AAAATTCCTTCACACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920
Db 4083 AAAATTCCTTCACACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTGGACTC 4142
QY |||||
QY 1921 AAGAACCCGCTCTCAGATCTCTCATCAAAAACAGCCCTGTTCTCGCAATCTCCGGCG 1980
Db 4143 AAGAACCCGCTCTCAGATCTCTCATCAAAAACAGCCCTGTTCTCGCAATCTCCGGCG 4202
QY |||||
QY 1981 GAGTTTTCAGCTAACAAAGTTTGTCTTCAATTCATCACCAATACTCCACAGGACAAGTGA 2040
Db 4203 GAGTTTTCAGCTAACAAAGTTTGTCTTCAATTCATCACCAATACTCCACAGGACAAGTGA 4262
QY |||||
QY 2041 GTGGAAATTTGATTTGAGGAGCTGCAGAAAGAAACAGCGCTGGAATCCCGAAGTGCAG 2100
Db 4263 GTGGAAATTTGATTTGAGGAGCTGCAGAAAGAAACAGCGCTGGAATCCCGAAGTGCAG 4322
QY |||||
QY 2101 TACACATCCAAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTT 2160
Db 4323 TACACATCCAAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTT 4382
QY |||||
QY 2161 TATACTGAGCTTCGCCCATTTGGCACCCGTTACCTTTACCTGCTCCCTGT 2209
Db 4383 TATACTGAGCTTCGCCCATTTGGCACCCGTTACCTTTACCTGCTCCCTGT 4431

RESULT 9

ADG39758

ID ADG39758 standard; DNA; 4718 BP.

XX ADG39758;

XX AC ADG39758;

XX DT 11-MAR-2004 (first entry)

XX AAV-1 genomic DNA sequence SEQ ID NO:20.

DE parvovirus; rep; cap; DNA binding domain; capsid interacting domain;

KW recombinant hybrid parvovirus particle;

KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;

KW gene; ds.

XX Adeno-associated virus 1.

XX OS

XX PN WO2003104392-A2.

XX XX

XX PD 18-DEC-2003.

XX XX

XX PF 02-DEC-2002; 2002WO-US038423.

XX XX

XX PR 18-DEC-2001; 2001US-0341919P.

XX XX

XX PA (UYNC-) UNIV NORTH CAROLINA.

661 GTGGGTAATGCTCAGGAAATTTGGCAATTCACATGCTGGGCGACAGAGTCATC 720
Db
5281 GTGGGTAATGCTCAGGAAATTTGGCAATTCACATGCTGGGCGACAGAGTCATC 5340
Qy
721 ACCACGAGCACCAGCAGCTGGGCTTTGCCACCTACAAATTAACCAACCTCTTACAAGCAAAATC 780
Db
5341 ACCACGAGCACCAGCAGCTGGGCTTTGCCACCTACAAATTAACCAACCTCTTACAAGCAAAATC 5400
Qy
781 TCAGTGTCTTAAAGGGGGCCAGCAAGCAACCACTACTTCGGCTCAGACACCCCTGG 840
Db
5401 TCAGTGTCTTAAAGGGGGCCAGCAAGCAACCACTACTTCGGCTCAGACACCCCTGG 5460
Qy
841 GGGTATTTTGAATTTCAACAGATTTCCACTGCGCACTTTTACACAGCTGAGCTGGCAGGACTC 900
Db
5461 GGGTATTTTGAATTTCAACAGATTTCCACTGCGCACTTTTACACAGCTGAGCTGGCAGGACTC 5520
Qy
901 ATCAACAAATTTGGGGAATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
Db
5521 ATCAACAAATTTGGGGAATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 5580
Qy
961 GTCAGGAGGTACAGACGAATGATGGCGTCAACCAATCGCTTAATACCTTACGACGAG 1020
Db
5581 GTCAGGAGGTACAGACGAATGATGGCGTCAACCAATCGCTTAATACCTTACGACGAG 5640
Qy
1021 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGCTCGGCTCTGGCACCAG 1080
Db
5641 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGCTCGGCTCTGGCACCAG 5700
Qy
1081 GGTGCTCTCCCTCCGTTCCGGCGGAGCTGTTTCATGATTCGGAATACCGCTACCTGAGC 1140
Db
5701 GGTGCTCTCCCTCCGTTCCGGCGGAGCTGTTTCATGATTCGGAATACCGCTACCTGAGC 5760
Qy
1141 CTCACAAATTTGGGCAAGCGTGGGACGTTTCATCTTTTACTGCTGGAAATATTTCCT 1200
Db
5761 CTCACAAATTTGGGCAAGCGTGGGACGTTTCATCTTTTACTGCTGGAAATATTTCCT 5820
Qy
1201 TCTCAGATCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
Db
5821 TCTCAGATCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 5880
Qy
1261 TTCCACAGCAGCTTACCGCACAGCCAGAGCCTGGACCGCTGATGTAATCTCTCATCGAC 1320
Db
5881 TTCCACAGCAGCTTACCGCACAGCCAGAGCCTGGACCGCTGATGTAATCTCTCATCGAC 5940
Qy
1321 CAATACCTGATTTACTGTAACAGAACTCAAAATCACTCGGAGTCCGCAAAACAAAGGAC 1380
Db
5941 CAATACCTGATTTACTGTAACAGAACTCAAAATCACTCGGAGTCCGCAAAACAAAGGAC 6000
Qy
1381 TTGCTGTTAGCCGTGGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTACCT 1440
Db
6001 TTGCTGTTAGCCGTGGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTACCT 6060
Qy
1441 GGACCTGTTTATCGGAGCAGCGCGTTTCTAAACAAACAAACAGACAAACAGCAAT 1500
Db
6061 GGACCTGTTTATCGGAGCAGCGCGTTTCTAAACAAACAAACAGACAAACAGCAAT 6120
Qy
1501 TTTACCTGGACTGGTGTCTTAAATATAACCTCAATGGGCGTGAATTCATCAACCT 1560
Db
6121 TTTACCTGGACTGGTGTCTTAAATATAACCTCAATGGGCGTGAATTCATCAACCT 6180
Qy
1561 GGCACTGTCTATGGCTTACAAAGACGAGCAAGTCTTCTTCCCATGAGCGGTGC 1620
Db
6181 GGCACTGTCTATGGCTTACAAAGACGAGCAAGTCTTCTTCCCATGAGCGGTGC 6240
Qy
1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCAATTTGGACAAATGTCATGATT 1680
Db
6241 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCAATTTGGACAAATGTCATGATT 6300
Qy
1681 ACAGACGAAGAGAAATTAAGCCACTTAACCTCTGTGGCCACCAAGAAAGATTGGGACCGTG 1740
Db
6301 ACAGACGAAGAGAAATTAAGCCACTTAACCTCTGTGGCCACCAAGAAAGATTGGGACCGTG 6360
Qy
1741 GCAGTCAATTTCCAGAGCAGCAGACAGACCCCTGCGACCGAGATGTGCTATGGGA 1800

6361 GCAGTCAATTTCCAGAGCAGCAGACAGACCTTCGGACCGAGATGTGCTATGGGA 6420
Qy
1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTTACTTCGACGGTCCCATTTGGGCC 1860
Db
6421 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTTACTTCGACGGTCCCATTTGGGCC 6480
Qy
1861 AAAATTTCTCACAGATGGACATTTTACCCGCTCTCTTATGGGCGGCTTTGGACTC 1920
Db
6481 AAAATTTCTCACAGATGGACATTTTACCCGCTCTCTTATGGGCGGCTTTGGACTC 6540
Qy
1921 AAGAACCCGCTCTCTCACAGATCTCTATCAAAAACACGCTGTTCTTGCAGAACTCTCCGGCG 1980
Db
6541 AAGAACCCGCTCTCTCACAGATCTCTATCAAAAACACGCTGTTCTTGCAGAACTCTCCGGCG 6600
Qy
1981 GAGTTTTCAGCTACAAAGTTTGTCTTATCATACCCCAATATCTCACAGGCAAGTGAAT 2040
Db
6601 GAGTTTTCAGCTACAAAGTTTGTCTTATCATACCCCAATATCTCACAGGCAAGTGAAT 6660
Qy
2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAG 2100
Db
6661 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAG 6720
Qy
2101 TACACATCCAAATTTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTT 2160
Db
6721 TACACATCCAAATTTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGACTT 6780
Qy
2161 TATACAGCTCGCCCAATTTGGCAACCGTTTACCTTACCCGCTCCCTGT 2209
Db
6781 TATACAGCTCGCCCAATTTGGCAACCGTTTACCTTACCCGCTCCCTGT 6829

RESULT 11

ADZ26928

ID ADZ26928 standard; DNA; 2211 BP.

XX

AC ADZ26928;

XX

XX 30-JUN-2005 (first entry)

XX

DE Adeno-associated virus DNA SEQ ID NO 78.

XX

KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antineumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

XX

OS Adeno-associated virus.

XX

PN WO2005033321-A2.

XX

PD 14-APR-2005.

XX

PF 30-SEP-2004; 2004WO-US028817.

XX

PR 30-SEP-2003; 2003US-0508226P.

XX

PR 29-APR-2004; 2004US-0566546P.

XX

PA (UYPE-) UNIV PENNSYLVANIA.

XX

XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX

DR WPI; 2005-285437/29.

XX

PT New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.

XX

PS Claim 19; SEQ ID NO 78; 569pp; English.

XX

CC The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX

SQ Sequence 2211 BP; 575 A; 654 C; 541 G; 441 T; 0 U; 0 Other;

Query Match 99.4%; Score 2196.2; DB 14; Length 2211;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2201; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCGCATTCGC	60
DB	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCGCATTCGC	60
QY	61	GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCCAACAGCAAAAGCAGGAC	120
DB	61	GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCGGAGCCCAAGCCCAACAGCAAAAGCAGGAC	120
QY	121	GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAGTACTCGGACCTTCAACGACTCGAC	180
DB	121	GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAGTACTCGGACCTTCAACGACTCGAC	180
QY	181	AAGGGGAGCCGCTCAACCGCGCGGACGACGGCCCTCGAGCAACAGGCGCTACGAC	240
DB	181	AAGGGGAGCCGCTCAACCGCGCGGACGACGGCCCTCGAGCAACAGGCGCTACGAC	240
QY	241	CAGCAGCTCAAGCGGGTGACATCCGTACTCGGTATTAACCGCGGACCGGATTT	300
DB	241	CAGCAGCTCAAGCGGGTGACATCCGTACTCGGTATTAACCGCGGACCGGATTT	300
QY	301	CAGGAGCGTCTGCAAGAAGTACGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG	360
DB	301	CAGGAGCGTCTGCAAGAAGTACGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG	360
QY	361	GCCAAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGCTCCT	420
DB	361	GCCAAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGCTCCT	420
QY	421	GGAAAGAACGTCGGTAGAGCAGTCCGCAACAGAGCCAGACTCTCTCGGGCATCGGC	480
DB	421	GGAAAGAACGTCGGTAGAGCAGTCCGCAACAGAGCCAGACTCTCTCGGGCATCGGC	480
QY	481	AAGACAGGCCAGCGCCGCTAAAGAGAGACTCAATTTTGGTCCAGACTGGCGACTCAGAG	540
DB	481	AAGACAGGCCAGCGCCGCTAAAGAGAGACTCAATTTTGGTCCAGACTGGCGACTCAGAG	540
QY	541	TCAGTCCCCGATCCAAACCTCTCGGAGAACTCTCCAGCAACCCCGCTGTGTGGACCT	600
DB	541	TCAGTCCCCGATCCAAACCTCTCGGAGAACTCTCCAGCAACCCCGCTGTGTGGACCT	600
QY	601	ACTACAAATGGCTTCAGCGGTGGGCAACCAATGGCGAGCAATTAACGAGCGCGGAG	660
DB	601	ACTACAAATGGCTTCAGCGGTGGGCAACCAATGGCGAGCAATTAACGAGCGCGGAG	660
QY	661	GTGGTTAATGCCCTCAGGAAATGGCATTTGGATTCCACATGGCTGGGCGACAGTATC	720
DB	661	GTGGTTAATGCCCTCAGGAAATGGCATTTGGATTCCACATGGCTGGGCGACAGTATC	720
QY	721	ACCACAGCACCCGACCTGGGCTTGGCCACCTTACAATAACCACTCTTACAAGCAATC	780
DB	721	ACCACAGCACCCGACCTGGGCTTGGCCACCTTACAATAACCACTCTTACAAGCAATC	780
QY	781	TCCAGTGTCTCAACGGGGGCGAGCAACCACTACTTCTGGGCTACAGCACCCCTGG	840
DB	781	TCCAGTGTCTCAACGGGGGCGAGCAACCACTACTTCTGGGCTACAGCACCCCTGG	840

QY	841	GGGTATTTTGATTTTCAACAGATTCCACTGCGCACTTTTCAACAGTGACTGGCAGGACTC	900
DB	841	GGGTATTTTGATTTTCAACAGATTCCACTGCGCACTTTTCAACAGTGACTGGCAGGACTC	900
QY	901	ATCAACAACATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA	960
DB	901	ATCAACAACATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA	960
QY	961	GTCAAGGAGGTCAACGCAATGATGGCGTCAACCATCGTCTTAACACCTTACACGACG	1020
DB	961	GTCAAGGAGGTCAACGCAATGATGGCGTCAACCATCGTCTTAACACCTTACACGACG	1020
QY	1021	GTTCAGTCTTCTCGGACTCGGAGTACAGACTTCCGTAGTCTCGGCTCTCGGACACG	1080
DB	1021	GTTCAGTCTTCTCGGACTCGGAGTACAGACTTCCGTAGTCTCGGCTCTCGGACACG	1080
QY	1081	GGTCCCTCCCTCCGTTCCGGCGGACGTTTCATGATTCGCAATACGCTACCTGACG	1140
DB	1081	GGTCCCTCCCTCCGTTCCGGCGGACGTTTCATGATTCGCAATACGCTACCTGACG	1140
QY	1141	CTCAACAATGGCAGCAAGCCCGTGGGACGTTTCATCTTTTACTGCTTGGAAATATTTCCCT	1200
DB	1141	CTCAACAATGGCAGCAAGCCCGTGGGACGTTTCATCTTTTACTGCTTGGAAATATTTCCCT	1200
QY	1201	TCTCAGATGCTCAGAAAGGGGCAACATTTTACCTTCAGCTACACCTTTGAGGAAGTCCCT	1260
DB	1201	TCTCAGATGCTCAGAAAGGGGCAACATTTTACCTTCAGCTACACCTTTGAGGAAGTCCCT	1260
QY	1261	TTTCCACAGGCTACGCGCACAGCAGGCTGGACCGGCTGATGAATCTCTCATCGAC	1320
DB	1261	TTTCCACAGGCTACGCGCACAGCAGGCTGGACCGGCTGATGAATCTCTCATCGAC	1320
QY	1321	CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAAGTCCGGAAGTCCCAAAACAAAGGAC	1380
DB	1321	CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAAGTCCGGAAGTCCCAAAACAAAGGAC	1380
QY	1381	TTTGTCTTTAGCGTGGGTCTCAGCTGGCATGTCTTTCAGCCCAAAACTGGCTACCT	1440
DB	1381	TTTGTCTTTAGCGTGGGTCTCAGCTGGCATGTCTTTCAGCCCAAAACTGGCTACCT	1440
QY	1441	GGACCTGTATTCGGCAGCAGCGCTTCTTAAACAAACAAACAGACACAAACAAACAAAT	1500
DB	1441	GGACCTGTATTCGGCAGCAGCGCTTCTTAAACAAACAAACAGACACAAACAAACAAAT	1500
QY	1501	TTTACTGTGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAACCCCT	1560
DB	1501	TTTACTGTGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAACCCCT	1560
QY	1561	GGCAGTGTATGGCTTCAACAAAGACGAGCAAGTCTTTTCCATGAGCGGTGTC	1620
DB	1561	GGCAGTGTATGGCTTCAACAAAGACGAGCAAGTCTTTTCCATGAGCGGTGTC	1620
QY	1621	ATGATTTTGGAAAGAGAGCGCGGACTTCAACACTGCAATGGCAATGTTCATGATT	1680
DB	1621	ATGATTTTGGAAAGAGAGCGCGGACTTCAACACTGCAATGGCAATGTTCATGATT	1680
QY	1681	ACAGAGCAAGAGAAATTAAGCCACTAACCCTCAATGGCGGTGAATCCATCATACCT	1740
DB	1681	ACAGAGCAAGAGAAATTAAGCCACTAACCCTCAATGGCGGTGAATCCATCATACCT	1740
QY	1741	GCAGTCAATTTCCAGAGCAGCAGACAGACCTTGGCCACCGAAAGATTTGGGACCGTG	1800
DB	1741	GCAGTCAATTTCCAGAGCAGCAGACAGACCTTGGCCACCGAAAGATTTGGGACCGTG	1800
QY	1801	GATTAACCTGGCATGGTGGCAAGATAGAGAGTGTACTGCGAGGTTCCCATTTGGGCC	1860
DB	1801	GATTAACCTGGCATGGTGGCAAGATAGAGAGTGTACTGCGAGGTTCCCATTTGGGCC	1860
QY	1861	AAATTTCTTCACACAGATGGACATTTTCAACCGGCTCTTATGGCGGCTTTGACTC	1920
DB	1861	AAATTTCTTCACACAGATGGACATTTTCAACCGGCTCTTATGGCGGCTTTGACTC	1920
QY	1921	AAGAACCCGCTCTCTCAGATCTCTCATCAAAACACAGCTGTCTTCTCGGAATCTCCGGCG	1980

```

Db 1921 AAGAACCGCCCTCTCAGATCTCATCAAAACACGCGCTGTTCTCGGAATCTCTCGGCG 1980
QY 1981 GAGTTTTCAGCTACAAAGTTTCTTCATTCATCAACCCATATCTCAACAGGACAAAGTGAGT 2040
Db 1981 GAGTTTTCAGCTACAAAGTTTCTTCATTCATTCATCAACCCATATCTCAACAGGACAAAGTGAGT 2040
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACACGCGCTGGAAATCCCGAAGTGCAG 2100
Db 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACACGCGCTGGAAATCCCGAAGTGCAG 2100
QY 2101 TACACATCAATATGCAAAATCTGCCAACGTTGATTTTACTGTGTGGACAAACAATGGACTT 2160
Db 2101 TACACATCAATATGCAAAATCTGCCAACGTTGATTTTACTGTGTGGACAAACAATGGACTT 2160
QY 2161 TATACAGCCTCGCCCAATTTGGACCCCGTTACCTTACCCGTCCTCGT 2209
Db 2161 TATACAGCCTCGCCCAATTTGGACCCCGTTACCTTACCCGTCCTCGT 2209

RESULT 12
AD226930
ID AD226930 standard; DNA; 2214 BP.
XX
AC AD226930;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus DNA SEQ ID NO 80.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antineumatic;
KW antihistatic; neuroprotective; antiinflammatory; antidiabetic;
KW antiposrotic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW viricide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
PN WO200503321-A2.
XX
PD 14-APR-2005.
XX
PP 30-SEP-2004; 2004WO-US028817.
XX
PR 30-SEP-2003; 2003US-0508226P.
XX 29-APR-2004; 2004US-056546P.
XX (UTPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
XX members, useful for preventing and/or treating arthritis, multiple
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
XX infection and cancer.
XX
XX Claim 50; SEQ ID NO 80; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
XX at least three AAV members, where each member of the AAV clade is
XX phylogenetically related as determined using a Neighbor-Joining heuristic
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
XX correction distance measurement of no more than 0.05. The methods and
XX compositions of the present invention are useful for the prevention
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
XX colitis. The present sequence represents an adeno-associated virus DNA.
XX
XX Sequence 2214 BP; 572 A; 658 C; 543 G; 441 T; 0 U; 0 Other;

```

```

Query Match 96.2%; Score 2124; DB 14; Length 2214;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

QY 1 ATGGCTGCGGATGGTTATCTTCCAGATTTGGCTCGAGGACAACTCTCTGAGGGCATTTCCG 60
Db 1 ATGGCTGCTGACGGTTATCTTCCAGATTTGGCTCGAGGACAACTCTCTGAGGGCATTTCCG 60
QY 61 GAGTGGTGGGACTTTGAAACCTTGAGGCCCGGAAAGCCCAACCAAGCCAAAGCAAGAGCAGAC 120
Db 61 GAGTGGTGGGACTTTGAAACCTTGAGGCCCGGAAAGCCCAACCAAGCCCAAGCAGAGCAGAC 120
QY 121 GACGGCGGGGCTGTGGTCTTCTGGCTTCAAGTACCTCGGACCTTCAACGGGCTCCGAC 180
Db 121 GACGGCGGGGCTGTGGTCTTCTGGCTTCAAGTACCTCGGACCTTCAACGGGCTCCGAC 180
QY 181 AAGGGGAGAGCCGTCACGCGCGGACGACGCGGCTTCGAGACGACCAAGGCTTACGAC 240
Db 181 AAGGGGAGAGCCGTCACGCGCGGACGACGCGGCTTCGAGACGACCAAGGCTTACGAC 240
QY 241 CAGCAGCTCAAGCGGGTGACAATCCGTTACTGGGTATTAACCAAGCCGACGCGGAGTTT 300
Db 241 CAGCAGCTCAAGCGGGTGACAATCCGTTACTGGGTATTAACCAAGCCGACGCGGAGTTT 300
QY 301 CAGGAGCGTCTCAAGAGATACGTTCTTTGGGGGCACTCTGGGCGGAGGCTTCTCCAG 360
Db 301 CAGGAGCGTCTCAAGAGATACGTTCTTTGGGGGCACTCTGGGCGGAGGCTTCTCCAG 360
QY 361 GCCAAGAAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCCT 420
Db 361 GCCAAGAAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCCT 420
QY 421 GGAAGAAACGTCGGGTAGACGTCGCCACA---AGAGCCAGACTCTCTCTCGGGCATC 477
Db 421 GGAAGAAAGACGCGGTAGAACCGTCACTCAGCGTTCTCCCGCATCTCTCAACGGGCATC 480
QY 478 GCAAGACAGGCGGACGCGCTTAAAGAGACTCAATTTTGGTCAGACTGGGCACTCA 537
Db 481 GCAAGAAAGCGGACGCGCTTAAAGAGACTCAATTTTGGTCAGACTGGGCACTCA 540
QY 538 GAGTCAGTCCCGCATCCAACTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGA 597
Db 541 GAGTCAGTCCCGCATCCAACTCTCGGAGAACCAACCAAGGCGCTCTGTTGTGGGA 600
QY 598 CTTACTCAATAGGCTTCAGGCGGTGGCGCAACAATGGCAGACAATAAGAGGCGCCGAC 657
Db 601 TCTGGTACAATGGCTGCAGGCGGTGGCGCTCCAATGGCAGACAATAAGAGGCGCCGAC 660
QY 658 GGAGTGGGTAAATGCTCAGGAAATTTGGCAATTTGGCAATTTGGCAATTTGGCAATTTGG 717
Db 661 GGAGTGGGTAAATGCTCAGGAAATTTGGCAATTTGGCAATTTGGCAATTTGGCAATTTGG 720
QY 718 ATCCACACAGCAGCCGCACTCGGCGCTTGGCCACTTACCAATTAACCACTCTTACAAGCAA 777
Db 721 ATCCACACAGCAGCCGCACTCGGCGCTTGGCCACTTACCAATTAACCACTCTTACAAGCAA 780
QY 778 ATCTCCAGTGTCTCAACGGGGGCCAGCAACGCAACCACTACTTCTGGGTACAGCACCCCC 837
Db 781 ATCTCCAGTGTCTTCAACGGGGGCCAGCAACGCAACCACTACTTCTGGGTACAGCACCCCC 840
QY 838 TGGGGGTATTTGATTTTCAACAGATTTCCAGTCCCACTTTTCCACCGTGAAGTGGCAGCA 897
Db 841 TGGGGGTATTTGATTTTCAACAGATTTCCAGTCCCACTTTTCCACCGTGAAGTGGCAGCA 900
QY 898 CTCATCAACAACTTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC 957
Db 901 CTCATCAACAACTTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC 960
QY 958 CAAAGTCAAGAGGTCACGACGAATGATGGCGTCAACCAATCGCTTAATTAACCTTACCAGC 1017
Db 961 CAAAGTCAAGAGGTCACGACGAATGATGGCGTCAACCAATCGCTTAATTAACCTTACCAGC 1020

```

QY 1018 ACGGTTCAAGTCTTTCTCGAGCTCGAGTACCAAGCTTCCGTAAGTCTCTCGGCTCTGGCAG 1077
Db 1021 ACGGTTCAAGTCTTTCTCGAGCTCGAGTACCAAGCTTCCGTAAGTCTCTCGGCTCTGGCAG 1080
QY 1078 CAGGGTGGCTCCCTCCGTTCCGGGGAGTCTTCATGATTCGGAATACGGTACCTG 1137
Db 1081 CAGGGTGGCTCCCTCCGTTCCGGGGAGTCTTCATGATTCGGAATACGGTACCTG 1140
QY 1138 ACGCTCAACAATGGCAGCCGAGCGTGGGACGTTTCATCTCTTTTACTGCTTGGAAATATTC 1197
Db 1141 ACGCTCAACAATGGCAGCCGAGCGTGGGACGTTTCATCTCTTTTACTGCTTGGAAATATTC 1200
QY 1198 CTTTCTCAGATGCTGAGAACGGGCAACAATTTTACCTTTCAGTACACCTTTGAGGAAGTG 1257
Db 1201 CTTTCTCAGATGCTGAGAACGGGCAACAATTTTACCTTTCAGTACACCTTTGAGGAAGTG 1260
QY 1258 CTTTCTCAGAGCTACCGGCACAGCCAGAGCTGGACCGGCTGATGATCTCTCATC 1317
Db 1261 CTTTCTCAGAGCTACCGGCACAGCCAGAGCTGGACCGGCTGATGATCTCTCATC 1320
QY 1318 GACCAATACCTGATTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAG 1377
Db 1321 GTCCAATACCTGATTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAG 1380
QY 1378 GACTTGTCTTTAGCCGTGGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTA 1437
Db 1381 GACTTGTCTTTAGCCGTGGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTA 1440
QY 1438 CTTGGACCTGTTATCGGCAGCAGCGGTTTCTAAACAAACAGACACACACACAC 1497
Db 1441 CTTGGACCTGTTATCGGCAGCAGCGGTTTCTAAACAAACAGACACACACACAC 1500
QY 1498 AATTTTACCTGGACTGGTGTCTCAAAATATAAATCAATCGGCGTGAATCCATCATCAAC 1557
Db 1501 AATTTTACCTGGACTGGTGTCTCAAAATATAAATCAATCGGCGTGAATCCATCATCAAC 1560
QY 1558 CTTGGCACTGCTATGGCTCACAAAGACGACGAAAGCAAGTCTTTTCCATGAGCGGT 1617
Db 1561 CTTGGCACTGCTATGGCTCACAAAGACGACGAAAGCAAGTCTTTTCCATGAGCGGT 1620
QY 1618 GTCATGATTTTGGAAAAGAGCGCGGAGCTTCAAACTGCGATTTGGACATGTCATG 1677
Db 1621 GTCATGATTTTGGAAAAGAGCGCGGAGCTTCAAACTGCGATTTGGACATGTCATG 1680
QY 1678 ATTACAGCAAGAGGAAATTAAGCCACTAACCTGTGCGCCACGAAAGATTGGGACC 1737
Db 1681 ATTACAGCAAGAGGAAATTAAGCCACTAACCTGTGCGCCACGAAAGATTGGGACC 1740
QY 1738 GTGGCAGTCAATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCATGCTATG 1797
Db 1741 GTGGCAGTCAATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCATGCTATG 1800
QY 1798 GGAGCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGAGGCTCCCATTTGG 1857
Db 1801 GGAGCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGAGGCTCCCATTTGG 1860
QY 1858 GCCAAATTCCTCACACAGATGACACTTTTACCCCGTCTCTCTTATGGCGGCTTTGGA 1917
Db 1861 GCCAAATTCCTCACACAGATGACACTTTTACCCCGTCTCTCTTATGGCGGCTTTGGA 1920
QY 1918 CTCAGAACCCGCTCTCTCAGATCTCTCATCAAAACACGCGTGTCTCGGAATCTCTCCG 1977
Db 1921 CTCAGAACCCGCTCTCTCAGATCTCTCATCAAAACACGCGTGTCTCGGAATCTCTCCG 1980
QY 1978 GCGAGATTTTTCAGTCAAAAGTTTGTCTTATTCATACCCCAATCTCCAGGACAGTG 2037
Db 1981 GCGAGATTTTTCAGTCAAAAGTTTGTCTTATTCATTCACCCCAATCTCCAGGACAGTG 2040
QY 2038 AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGACGCGCTGGAATCCCGAAGTG 2097
Db 2041 AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGACGCGCTGGAATCCCGAAGTG 2100
QY 2098 CAGTACACATCTCAATTTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGA 2157

Db 2101 CAGTACACATCCAAATATGCAAAATCTGCCAGCGTTGATTTACTGTGACAAACAATGGA 2160
QY 2158 CTTTATACTAGCCTCGCCCCATTTGGCACCCGTTTACCTTTACCGTCCCTGT 2209
Db 2161 CTTTATACTAGCCTCGCCCCATTTGGCACCCGTTTACCTTTACCGTCCCTGT 2212
RESULT 13
ADZ26931
ID ADZ26931 standard; DNA; 2211 BP.
XX
AC ADZ26931;
XX
DT 30-JUN-2005 (first entry)
XX
DE Adeno-associated virus DNA SEQ ID NO 81.
XX
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes; HIV;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antihistaminic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
PN WO200503321-A2.
XX
PD 14-APR-2005.
XX
PF 30-SEP-2004; 2004WO-US029817.
XX
PR 30-SEP-2003; 2003US-0508226P.
XX
PS 29-APR-2004; 2004US-0566546P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;
XX
PS WPI; 2005-285437/29.
XX
PT New adeno-associated virus (AAV) clade comprising at least three AAV
members, useful for preventing and/or treating arthritis, multiple
sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
infection and cancer.
XX
PS Claim 19; SEQ ID NO 81; 569pp; English.
XX
CC The invention relates to an adeno-associated virus (AAV) clade comprising
at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX
SQ Sequence 2211 BP; 582 A; 653 C; 535 G; 441 T; 0 U; 0 Other;
Query Match 95.5%; Score 2109.8; DB 14; Length 2211;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGACAACTCTCTGAGGCAATTCG 60
Db 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGACACTCTCTCTGAAGAAATAAGA 60
QY 61 GAGTGTGGGACTTGAACCTGGAGCCCCGAAACCCCAACCCAGCAACAGCAAGAGCGGAC 120
Db 61 CAGTGTGGGAGCTCGAGACTGGCCCCACCACCACCAAGCCCGAGAGCGGCAATAGGAC 120

121 GACGGCGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGACCTCGAC 180
121 GACAGCGGGTCTTGTGCTTCTGGGTACAAGTACCTCGGACCTTCAACGACCTCGAC 180
181 AAGGGGAGCCGCTCAACCGGCGGACGAGCGGCGCTCGAGCAGACAGGCGCTACGAC 240
181 AAGGAGAGCGGTCAACGAGGACGAGCGCGGCGCTCGAGCAGACAGGCGCTACGAC 240
241 CAGCAGCTCAAGCGGGTGACAAATCCGTACTCGGTATAAACGCGCAGCGCGAGTTT 300
241 CGGACGCTCGACAGCGGAGACAACCGGTACTCAAGTACAACCAACGCGCAGCGGAGTTT 300
301 CAGGAGCGTCTGCAAGAGATAGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
301 CAGGAGCGCTTAAAGAAATAGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
361 GCCAAGAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 420
361 GCCAAGAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTGAGACGGCTCCT 420
421 GGAAGAAAGCTCCGGTGAAGCAGTCCGCCAAGAGCCAGACTCTCTCGGGCAGTCGGC 480
421 GGAAGAAAGCTCCGGTGAAGCAGTCCGCCAAGAGCCAGACTCTCTCGGGCAGTCGGC 480
481 AAGACAGCCAGCGCCCTTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540
481 AAGACAGCCAGCGCCCTTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540
541 TCAGTCCCGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGGGACCT 600
541 TCAGTCCCGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGGGACCT 600
601 ACTAATAGCTTCAGCGGTGGCGCAACCAATGGCAGACAATAAACGAGCGCGCAGGA 660
601 ACTAATAGCTTCAGCGGTGGCGCAACCAATGGCAGACAATAAACGAGCGCGCAGGA 660
661 GTGGTAAATGCTTCAGGAAATGGCAATGGCATTCGATTCACATGGCTGGCGCAGAGTCATC 720
661 GTGGTAAATGCTTCAGGAAATGGCAATGGCATTCGATTCACATGGCTGGCGCAGAGTCATC 720
721 ACCACAGACCCCGACCTGGGCGCTTGGCCACCTACAATAACCACTCTACAAGCAATC 780
721 ACCACAGACCCCGACCTGGGCGCTTGGCCACCTACAATAACCACTCTACAAGCAATC 780
781 TCAGTGTCTTCAACGGGGCGCAGCAACCACTACTTTCGGCTACAGCAACCCCGCTGG 840
781 TCAGTGTCTTCAACGGGGCGCAGCAACCACTACTTTCGGCTACAGCAACCCCGCTGG 840
841 GGGTATTTTGAATTCACAGATTCACCTGCCACTTTTCCACGACTGGCGAGCGACTC 900
841 GGGTATTTTGAATTCACAGATTCACCTGCCACTTTTCCACGACTGGCGAGCGACTC 900
901 ATCAACAAATTTGGGATTCGGGCGCAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960
901 ATCAACAAATTTGGGATTCGGGCGCAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960
961 GTCAAGAGGTCAAGCAATGATGGCGTCAACCAATCGCTAATAACCTTTACCAGCAG 1020
961 GTCAAGAGGTCAAGCAATGATGGCGTCAACCAATCGCTAATAACCTTTACCAGCAG 1020
1021 GTTCAAGTCTTTCGGACTCGGAGTACCAGTTCGGTACGTCTCGGCTCTGGGACCGAG 1080
1021 GTTCAAGTCTTTCGGACTCGGAGTACCAGTTCGGTACGTCTCGGCTCTGGGACCGAG 1080
1081 GGTGCTCCCTCCGTTCCCGGGGAGCGTTCATGATTCGCAATACGGCTACCTGAGC 1140
1081 GGTGCTCCCTCCGTTCCCGGGGAGCGTTCATGATTCGCAATACGGCTACCTGAGC 1140
1141 CTCAACAAATGGCAGCCAGCGGTGGGACGTTTCATCTCTTTTACTGCTGGGAATTTCCCT 1200
1141 CTCAACAAATGGCAGCCAGCGGTGGGACGTTTCATCTCTTTTACTGCTGGGAATTTCCCT 1200

RESULT 14
AD227053
ID AD227053 standard; DNA; 2211 BP.
XX

1201 TCTCAGATGCTGAGAACGGGCAACAACTTTACCTTACGCTACACCTTTGAGGAAGTGCT 1260
1201 TCTCAGATGCTGAGAACGGGCAACAACTTTACCTTACGCTACACCTTTGAGGAAGTGCT 1260
1261 TTCCAACAGCAGTACGGCGCACAGCGAGCCTTGGACCGGCTGATGAATCTCTCATCGAC 1320
1261 TTCCAACAGCAGTACGGCGCACAGCGAGCCTTGGACCGGCTGATGAATCTCTCATCGAC 1320
1321 CAATACCTGTATTTACCTGAACAGAACTCAAAATCAAGTCGGGAAGTCCCAAAACAAGAC 1380
1321 CAATACCTGTATTTACCTGAACAGAACTCAAAATCAAGTCGGGAAGTCCCAAAACAAGAC 1380
1381 TTGCTGTTTTAGCGTGGGTCTCCAGCTGGCATGCTGTTCCAGCCCAAAACTGGGTACT 1440
1381 TTGCTGTTTTAGCGTGGGTCTCCAGCTGGCATGCTGTTCCAGCCCAAAACTGGGTACT 1440
1441 GGACCTCTGTATTCGGCAGCAGCGGCTTTCTAAAACAAAAAACAGACAACCAACAGCAAT 1500
1441 GGACCTCTGTATTCGGCAGCAGCGGCTTTCTAAAACAAAAAACAGACAACCAACAGCAAT 1500
1501 TTTTACCTGGAAGTGGTCTTCAAAATATAAATCAATGGGCGTGAATCCATCATCAACCT 1560
1501 TTTTACCTGGAAGTGGTCTTCAAAATATAAATCAATGGGCGTGAATCCATCATCAACCT 1560
1561 GGCAGCTGCTATGGCTTCACAAAGACGACGAAAGAAAGTTCTTTCCCATGAGCGGTGC 1620
1561 GGCAGCTGCTATGGCTTCACAAAGACGACGAAAGAAAGTTCTTTCCCATGAGCGGTGC 1620
1621 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGACAATGTCTATGATT 1680
1621 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGACAATGTCTATGATT 1680
1681 ACAGACGAGAGAGAAATTAAGGCACTAAACCTGTGGGCAACGAAAGATTTGGGACCGTG 1740
1681 ACAGACGAGAGAGAAATTAAGGCACTAAACCTGTGGGCAACGAAAGATTTGGGACCGTG 1740
1741 GCAGTCAATTTCCAGACGACGACAGACACCTCGACCGGAGATGTCATGCTATGGGA 1800
1741 GCAGTCAATTTCCAGACGACGACAGACACCTCGACCGGAGATGTCATGCTATGGGA 1800
1801 GCATTTACCTTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGCGAGGTCCTTGGGCC 1860
1801 GCATTTACCTTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGCGAGGTCCTTGGGCC 1860
1861 AAAATTCCTCACACAGATGGACATTTTCAACCGTCTCTCTTATGGGCGGCTTTGGACTC 1920
1861 AAAATTCCTCACACAGATGGACATTTTCAACCGTCTCTCTTATGGGCGGCTTTGGACTC 1920
1921 AAGAACCCGCTCTCAGATCCTCATCAAAACACGCTGTTCCTGCGAATCCTCCGGCG 1980
1921 AAGAACCCGCTCTCAGATCCTCATCAAAACACGCTGTTCCTGCGAATCCTCCGGCG 1980
1981 GAGTTTTTCAAGTACAAAGTTTGTCTTCAATCATCCCAATACTCCACAGGACAAGTGAGT 2040
1981 GAGTTTTTCAAGTACAAAGTTTGTCTTCAATCATCCCAATACTCCACAGGACAAGTGAGT 2040
2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTCAG 2100
2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTCAG 2100
2101 TACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACATGACTT 2160
2101 TACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACATGACTT 2160
2161 TATACTGAGCCTCGCCCCCATTTGGCACCCGTTACCTTTACCCGTCCTCTGT 2209
2161 TATACTGAGCCTCGCCCCCATTTGGCACCCGTTACCTTTACCCGTCCTCTGT 2209

AC AD227053;
XX 30-JUN-2005 (first entry)
XX Adeno-associated virus DNA SEQ ID NO 203.
DE
XX
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;
XX virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.
XX
OS Adeno-associated virus.
XX
XX WO2005033321-A2.
XX
XX 14-APR-2005.
XX
XX 30-SEP-2004; 2004WO-US028817.
XX
XX 30-SEP-2003; 2003US-0508226P.
XX 29-APR-2004; 2004US-0566546P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;
XX WPI; 2005-285437/29.
XX
XX New adeno-associated virus (AAV) clade comprising at least three AAV
PT members, useful for preventing and/or treating arthritis, multiple
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial
PT infection and cancer.
XX
XX Disclosure; SEQ ID NO 203; 569pp; English.
XX
XX The invention relates to an adeno-associated virus (AAV) clade comprising
CC at least three AAV members, where each member of the AAV clade is
CC phylogenetically related as determined using a Neighbor-Joining heuristic
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson
CC correction distance measurement of no more than 0.05. The methods and
CC compositions of the present invention are useful for the prevention
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative
CC colitis. The present sequence represents an adeno-associated virus DNA.
XX
XX Sequence 2211 BP; 576 A; 647 C; 547 G; 441 T; 0 U; 0 Other;
SQ
Query Match 95.3%; Score 2105; DB 14; Length 2211;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 2144; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 1 ATGGCTGCCGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCAATTCG 60
DB 1 ATGGCTGCCGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCAATTCG 60
QY 61 GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCAAACCAAGCCAAAGCAGGAC 120
DB 61 GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCGGAGCCCAAGCCAAACCAAGCAGGAC 120
QY 121 GAGCGCCGGGTCTGGTCTCTTGGCTACAAGTACTCTCGGACCTTCACGGACTCGAC 180
DB 121 GAGCGCCGGGTCTGGTCTCTTGGCTACAAGTACTCTCGGACCTTCACGGACTCGAC 180
QY 181 AAGGGGAGCCCGTCAACGGCGGAGCGAGCGGCCCTCGAGCAACGAAGGCGCTACGAC 240
DB 181 AAGGGGAGCCCGTCAACGGCGGAGCGAGCGGCCCTCGAGCAACGAAGGCGCTACGAC 240
QY 241 CAGCAGCTCAAGCGGGTGACAAATCCGTACTCCGGTATACCAACGCCGACGGCGAGTTT 300
DB 241 CAGCAGCTCAAGCGGGTGACAAATCCGTACTCCGGTATACCAACGCCGACGGCGAGTTT 300

QY 301 CAGGAGCGCTCTGCAAGAGATACGCTTTTGGGGCAACCTCGGCGAGCAGCTCTTCCAG 360
DB 301 CAGGAGCGCTCTGCAAGAGATACGCTTTTGGGGCAACCTCGGCGAGCAGCTCTTCCAG 360
QY 361 GCCAAGAAGCGGTTCTCGAACCTCTCGGCTGCTGGTTGAGGAGCGCTTAAGACGGCTCT 420
DB 361 GCCAAGAAGCGGTTCTCGAACCTTTTGGTCTGGTTGAGGAGGTTGCTTAAGACGGCTCT 420
QY 421 GGAAAGAAACGTCGCGGTAGAGCAGTCGCCAAGAAGCCAGACTCTCTCTCGGCGCATCGGC 480
DB 421 GGAAAGAAACGTCGCGGTAGAGCAGTCGCCAAGAAGCCAGACTCTCTCTCGGCGCATTGGC 480
QY 481 AAGACAGCCAGCAGCCCGCTTAAAGAGAGACTCAATTTTGGTTCAGACTGGCGCATCAGAG 540
DB 481 AAGACAGCCAGCAGCCCGCTTAAAGAGAGACTCAATTTTGGTTCAGACTGGCGCATCAGAG 540
QY 541 TCAGTCCCGATCCGACACCTCTCGGAGAACCTTCAGGAACCCCGCTGCTGTGGGACCT 600
DB 541 TCAGTCCCGACCCCAACCTCTCGGAGAACCTTCAGGAACCCCGCTGCTGTGGGACCT 600
QY 601 ACTACAAATGGCTTCAGGGCGGTGGCGCACCAATGGGAGACAATAAGAGCGCCGACGGA 660
DB 601 ACTACAAATGGCTTCAGGGCGGTGGCGCACCAATGGGAGACAATAAGAGCGCCGACGGA 660
QY 661 GTGGGTAATGCTCTCAGGAAATTTGGCATTTGCGATTCCACATGGCTGGCGCACAGTCTATC 720
DB 661 GTGGGTAATGCTCTCAGGAAATTTGGCATTTGCGATTCCACATGGCTGGCGCACAGTCTATC 720
QY 721 ACCACAGCAGCCGACCTGGGCGCTTGGCCACCTACATTAACCACTCTTACAGCAAAATC 780
DB 721 ACCACAGCAGCCGACCTGGGCGCTTGGCCACCTTAAACCACTCTTACAGCAAAATC 780
QY 781 TCCAGTGTCTCAACGGGGCCAGCAACGACCACTACTCTCGGCTACAGCAGCCCGCTGG 840
DB 781 TCCAGTGTCTCAACGGGGCCAGCAACGACCACTACTCTCGGCTACAGCAGCCCGCTGG 840
QY 841 GGGTATTTTGAATTTCAACAGATTCCACTGSCCACTTTTCAACAGTGACTGGCAGCGACTC 900
DB 841 GGGTATTTTGAATTTCAACAGATTCCACTGSCCACTTTTCAACAGTGACTGGCAGCGACTC 900
QY 901 ATCAACAACAAATGGGGATTCCGGCCCAAGAGACTCACTTCAACTCTTCAACATCCAA 960
DB 901 ATCAACAACAAATGGGGATTCCGGCCCAAGAGACTCACTTCAAGTCTTCAACATCCAA 960
QY 961 GTCAAGGAGGTCAACGAGCAATGATGCGCTCAACCACTCGCTAATAACCTTACGACAGC 1020
DB 961 GTCAAGGAGGTCAACGAGCAATGATGCGCTCAACCACTCGCTAATAACCTTACGACAGC 1020
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTCTGCGCACCGAC 1080
DB 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTCTGCGCACCGAC 1080
QY 1081 GGTGCTCTCCCTCGGTTCCGGCGGAGCGTTCATGATTTCCGCAATACGGCTACCTGACG 1140
DB 1081 GGTGCTCTCCCTCGGTTCCGGCGGAGCGTTCATGATTTCCGCGTACGGCTACCTGACG 1140
QY 1141 CTCAACAATGGGAGCCGAGCGTGGGAGCGTTCATCTTTTACTGCTCGAATAATTTCCCT 1200
DB 1141 CTCAACAATGGGAGCCGAGCGTGGGAGCGTTCATCTTTTACTGCTCGAATAATTTCCCA 1200
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 1260
DB 1201 TCGCAGATGCTGAGAACGGGCAATACTTTTACCTTCAGCTACACCTTCGAGGAGTGCCT 1260
QY 1261 TTCCACAGAGCTACCGGCAAGCGGAGCTGGACCGGCTGATGATCTCTCATCGAC 1320
DB 1261 TTCCACAGAGCTACCGGCAAGCGGAGCTGGACCGGCTGATGATCTCTCATCGAC 1320
QY 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGAGAGTCCCCAAAAACAAGGAC 1380
DB 1321 CAGTACCTGTATTACCTGAACAGAACTCAGAAATCAGTCCGAGAGTCCCCAAAAACAAGGAC 1380

QY 601 ACTACAATGGCTTCAGCGGTTGGCGCAACCAATGGCAGACAATAACGAAGCGCCGACGGA 660
Db 2621 ACTACAATGGCTTCAGCGGTTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGGA 2680
QY 661 GTGGGTAAATGCCTCAGGAATTTGGCAATTTGGATTTCCACATGGCTGGGCGACAGATCATC 720
Db 2681 GTGGGTAAATGCCTCAGGAATTTGGCAATTTGGATTTCCACATGGCTGGGCGACAGATCATC 2740
QY 721 ACCACAGACCCGCACTGGGCTTTGCCACCTACAATAAACCACTCTTACAAGCAAAATC 780
Db 2741 ACCACAGACCCGCACTGGGCTTTGCCACCTACAATAAACCACTCTTACAAGCAAAATC 2800
QY 781 TCCAGTGTCTTCAACGGGGGCCAGCAACGACAACTACTTCCGGCTACAGCAACCCCTGG 840
Db 2801 TCCAGTGTCTTCAACGGGGGCCAGCAACGACAACTACTTCCGGCTACAGCAACCCCTGG 2860
QY 841 GGGTATTTTGAATTTCAACAGATTTCACTGGCACTTTTACACAGCTGACTGGGAGGACATC 900
Db 2861 GGGTATTTTGAATTTCAACAGATTTCACTGGCACTTTTACACAGCTGACTGGGAGGACATC 2920
QY 901 ATCAACAACTTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
Db 2921 ATCAACAACTTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 2980
QY 961 GTCAAGGAGGTCAACGAGATGATGGCGTCACAAACCATCGCTAATAACCTTACCAAGCAG 1020
Db 2981 GTCAAGGAGGTCAACGAGATGATGGCGTCACAAACCATCGCTAATAACCTTACCAAGCAG 3040
QY 1021 GTTCAAGTCTTTCGAGCTCGGAGTACAGCTTTCGGTACGTCTCGGCTTCGGGACCAAG 1080
Db 3041 GTTCAAGTCTTTCGAGCTCGGAGTACAGCTTTCGGTACGTCTCGGCTTCGGGACCAAG 3100
QY 1081 GGTGCTCCCTCCGTTCCGGCGGAGCTGTCATGATTCGGCAATACGGCTACCTGACG 1140
Db 3101 GGTGCTCCCTCCGTTCCGGCGGAGCTGTCATGATTCGGCAATACGGCTACCTGACG 3160
QY 1141 CTCACAAATGGCAGCAAGCGTGGGACGTTTCACTCTTTACTGCTGGAATATTTCCCT 1200
Db 3161 CTCACAAATGGCAGCAAGCGTGGGACGTTTCACTCTTTACTGCTGGAATATTTCCCA 3220
QY 1201 TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
Db 3221 TCGCAGATGCTGAGAAACGGGCAATAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 3280
QY 1261 TTCCAAGAGCTACCGGCACAGCCAGAGCTTCGAGCGGTGATGATCTCTCATCGAC 1320
Db 3281 TTCCAAGAGCTACCGGCACAGCCAGAGCTTCGAGCGGTGATGATCTCTCATCGAC 3340
QY 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTTCGGAAAGTGCCTCAAAACAAAGGAC 1380
Db 3341 CAGTACCTGTATTACCTGAACAGAACTCAAAATCAGTTCGGAAAGTGCCTCAAAACAAAGGAC 3400
QY 1381 TTGCTGTTTAGCGTGGGCTTCAGCTGGCATGCTGTTTCAGCCCAAAACTGGGTACCT 1440
Db 3401 TTGCTGTTTAGCGTGGGCTTCAGCTGGCATGCTGTTTCAGCCCAAAACTGGGTACCT 3460
QY 1441 GGAACCTGTATTTCGGCAGAGCGGTTTCAAAACAAACAGACACACACACACACACAC 1500
Db 3461 GGAACCTGTATTTCGGCAGAGCGGTTTCAAAACAAACAGACACACACACACACACAC 3520
QY 1501 TTTTACCTGGAAGTGGTCTTCAAAATATACTCAATGGGCGGTGAATCCATCATCAACCCCT 1560
Db 3521 TTTTACCTGGAAGTGGTCTTCAAAATATACTCAATGGGCGGTGAATCCATCATCAACCCCT 3580
QY 1561 GGCACCTGCTATGGCTTCACAAAGAGCAGAAAGCAAGTTCCTTCCATGAGCGGTGTC 1620
Db 3581 GGCACCTGCTATGGCTTCACAAAGAGCAGAAAGCAAGTTCCTTCCATGAGCGGTGTC 3640
QY 1621 ATGATTTTGGAAAAAGAGAGCGGCGGAGCTTCAAACTGCTGATTTGGCAATGCTATGAT 1680
Db 3641 ATGATTTTGGAAAAAGAGAGCGGCGGAGCTTCAAACTGCTGATTTGGCAATGCTATGAT 3700
QY 1681 ACAGACGAAGGAAATTAAGCCACTAACCTCTGGCCACCGAAAGATTTGGGACCGTG 1740

Db 3701 ACAGACGAAGGAAATCAAAAGCCACTAACCCCGTGGCCACCGAAAGATTTGGGACTGTG 3760
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCTCTGGACCCGAGATGTGATGCTATGGGA 1800
Db 3761 GCAGTCAATTTCCAGAGCAGCAGCAGACCTCTGGACCCGAGATGTGATGCTATGGGA 3820
QY 1801 GCATTAACCTGGCATGGTGTGGCAAGATAGAGACGTGTACTGCAGGGTCCCATTGGGCG 1860
Db 3821 GCATTAACCTGGCATGGTGTGGCAAGATAGAGACGTGTACTGCAGGGTCCCATTGGGCG 3880
QY 1861 AAAATTCCTCACAGATGGACACTTTTACCCGCTCTCTTATATGGGCGGCTTTGACTC 1920
Db 3881 AAAATTCCTCACAGATGGACACTTTTACCCGCTCTCTTATATGGGCGGCTTTGACTC 3940
QY 1921 AAGAACCCGCTCTCAGATCCTCAATAAACAAGACGCTGTCTCTGCGAATCCTCCGGCG 1980
Db 3941 AAGAACCCGCTCTCAGATCCTCAATAAACAAGACGCTGTCTCTGCGAATCCTCCGGCG 4000
QY 1981 GAGTTTTCAGCTACAAAGTTTCTTCAATCATCAACCAATCTCCAAGGACAAAGTGAGT 2040
Db 4001 GAGTTTTCAGCTACAAAGTTTCTTCAATCATCAACCAATCTCCAAGGACAAAGTGAGT 4060
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100
Db 4061 GTGGAGATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 4120
QY 2101 TACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGAACAACATGACTT 2160
Db 4121 TATACATCTAATCTATGCAAAATCTGCCAACGTTGATTTTACTGTGGAACAACATGACTT 4180
QY 2161 TATACTGAGCTTCGCCCAATTTGGCACCCGTTACTTACCCGTCCTCTGT 2209
Db 4181 TATACTGAGCTTCGCCCAATTTGGCACCCGTTACTTACCCGTCCTCTGT 4229

Search completed: November 28, 2005, 19:55:43

Job time : 1039.63 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 8072.36 Seconds
(without alignments)
15555.198 Million cell updates/sec

Title: US-10-696-900-1_COPY_2223_4431
Perfect score: 2209
Sequence: 1 atggctgcgatgttatct.....ttacctaccgcctccctgt 2209

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5881141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.in.*
- 3: gb.env.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pr.*
- 9: gb.roi.*
- 10: gb.sts.*
- 11: gb.sy.*
- 12: gb.un.*
- 13: gb.vi.*
- 14: gb.htg.*
- 15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2209	100.0	2211	6	BD242771 Adeno-ass
2	2209	100.0	2211	6	CS073491 Sequence
3	2209	100.0	2211	6	CS073614 Sequence
4	2209	100.0	2211	6	AR562503 Sequence
5	2209	100.0	4718	6	BD242766 Adeno-ass
6	2209	100.0	4718	6	AR562498 Sequence
7	2209	100.0	4718	6	AX753251 Sequence
8	2209	100.0	4718	13	AF063497 Adeno-ass
9	2202.6	99.7	7447	6	AR527492 Sequence
10	2196.2	99.4	2211	6	CS073490 Sequence
11	2196.2	99.4	2211	13	AY530611 Adeno-ass
12	2124	96.2	2214	6	CS073492 Sequence
13	2124	96.2	2214	13	AY530606 Adeno-ass
14	2109.8	95.5	2211	6	CS073493 Sequence
15	2109.8	95.5	2211	13	AY530607 Adeno-ass
16	2105	95.3	2211	6	CS073615 Sequence
17	2105	95.3	4683	13	AF028704 Adeno-ass
18	2101.8	95.1	4683	6	BD242775 Adeno-ass

19	2101.8	95.1	4683	6	AR562507 Sequence
20	2100.2	95.1	2211	6	CS073494 Sequence
21	2100.2	95.1	2211	13	AY530609 Adeno-ass
22	1798	81.4	1800	6	BD242772 Adeno-ass
23	1798	81.4	1800	6	AR562504 Sequence
24	1639	74.2	2214	6	CS073452 Sequence
25	1639	74.2	2214	13	AY530567 Adeno-ass
26	1620	73.3	2214	6	CS073453 Sequence
27	1616.8	73.2	2211	13	AY242997 Non-human
28	1616.8	73.2	3122	6	AX753290 Sequence
29	1616.6	73.2	2214	6	CS073449 Sequence
30	1616.6	73.2	2214	13	AY530568 Adeno-ass
31	1615.2	73.1	3117	6	AX753288 Sequence
32	1615.2	73.1	3122	6	AX753287 Sequence
33	1613.6	73.0	3121	6	AX753289 Sequence
34	1611.8	73.0	2214	6	CS073625 Sequence
35	1611.8	73.0	4721	6	CQ972062 Sequence
36	1611.8	73.0	4721	6	CS073592 Sequence
37	1611.8	73.0	4721	6	AX753246 Sequence
38	1611.8	73.0	4721	13	AF513851 Adeno-ass
39	1607	72.7	2214	6	CS073447 Sequence
40	1607	72.7	2214	6	CS073431 Sequence
41	1605.4	72.7	2214	6	CS073448 Sequence
42	1605.4	72.7	2214	13	AY243020 Non-human
43	1605.4	72.7	2217	6	CS073454 Sequence
44	1605.4	72.7	3127	6	AX753260 Sequence
45	1603.8	72.6	2214	6	CS073444 Sequence

ALIGNMENTS

RESULT 1

BD242771

LOCUS

DEFINITION

ADENO-ASSOCIATED VIRUS SERUM TYPE 1 NUCLEIC ACID SEQUENCE, VECTOR

BD242771

ACCESSION

BD242771.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2211 bp DNA linear PAT 17-JUL-2003
Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same.

BD242771.1 GI:33052541
JP 2002529098-A/6.
unidentified
unidentified
unclassified.

1 (bases 1 to 2211)
Wilson,J.M. and Xiao,W.

Adeno-associated virus serum type 1 nucleic acid sequence, vector
and host cell containing the same
Patent: JP 2002529098-A 6 10-SEP-2002.

THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
OS AAV-1

PN JP 2002529098-A/6
PD 10-SEP-2002

PP 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114

PI JAMES M WILSON, WEIDONG XIAO
PC C12N15/09,A61K31/711,A61P43/00,A61P43/00,C12N1/15,C12N1/19, PC

C12N1/21,
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC

vector and host
CC cell containing the same

FT Key Location/Qualifiers
CDS (1)..(2208).

Location/Qualifiers
1..2211
/organism="unidentified"

/mol_type="genomic DNA"
/db_xref="taxon:32644"

source

ORIGIN

Query Match 100.0%; Score 2209; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
Db      2161 TATACTGAGCTCGCCCATGGACCCGTTACCTTACCGCTCCCTGT 2209
|||||
RESULT 2
CS073491
LOCUS      2211 bp      DNA      linear      PAT 05-MAY-2005
DEFINITION      Sequence 79 from Patent WO2005033321.
ACCESSION      CS073491
VERSION      CS073491.1 GI:63090470
KEYWORDS      .
SOURCE      unidentified
ORGANISM      unidentified
              unclassified.
REFERENCE
1. Wilson, J.M., Gao, G., Alvira, M.R. and Vandenberghe, L.H.
   Adeno-associated virus (aav) clades, sequences, vectors containing
   same, and uses therefor
   Patent: WO 2005033321-A 79 14-APR-2005;
   The Trustees of the University of Pennsylvania (US)
   Location/Qualifiers
     1..2211
       /organism="unidentified"
       /mol_type="unassigned DNA"
       /db_xref="taxon:32644"
       /note="new AAV serotype, clone rh.71"

ORIGIN
Query Match      100.0%; Score 2209; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGCGATTTCG 60
Db      1 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGCGATTTCG 60

QY      61 GAGTGTGGGACTTGAACCTCGAGCGCCCGAAGCCCAAGCCCAACGAGCAAAAGCAGGAC 120
Db      61 GAGTGTGGGACTTGAACCTCGAGCGCCCGAAGCCCAAGCCCAACGAGCAAAAGCAGGAC 120

QY      121 GACGGCGGGGTCTGGTCTTCTGGCTACAAGTACTCGGACCTTCAACGAGCTCGAC 180
Db      121 GACGGCGGGGTCTGGTCTTCTGGCTACAAGTACTCGGACCTTCAACGAGCTCGAC 180

QY      181 AAGGGGAGCCCTCAACGGCGGGAGCGAGCGCCCTCGAGCAGCAAGGCGCTACGAC 240
Db      181 AAGGGGAGCCCTCAACGGCGGGAGCGAGCGCCCTCGAGCAGCAAGGCGCTACGAC 240

QY      241 CAGCAGCTCAAGCGGGTGACAATCCGTACTCTCGGTATTAACCAACGCGGACGCGAGTTT 300
Db      241 CAGCAGCTCAAGCGGGTGACAATCCGTACTCTCGGTATTAACCAACGCGGACGCGAGTTT 300

QY      301 CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGGAGAGCAGTCTTCCAG 360
Db      301 CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGGAGAGCAGTCTTCCAG 360

QY      361 GCCAAGAGCGGGTCTCGAACCTCTCGTCTGGTTGAGGAAGCGCTAGACGGCTCCT 420
Db      361 GCCAAGAGCGGGTCTCGAACCTCTCGTCTGGTTGAGGAAGCGCTAGACGGCTCCT 420

QY      421 GGAAGAAACGTCCGGTAGAGCAGTCCGACAGAGCGAGACTCTCTCGGGCATCGGC 480
Db      421 GGAAGAAACGTCCGGTAGAGCAGTCCGACAGAGCGAGACTCTCTCGGGCATCGGC 480

QY      481 AAGACAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTCGCGACTCAGAG 540
Db      481 AAGACAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTCGCGACTCAGAG 540

QY      541 TCAGTCCCCGATCCACAACCTCTCGGAGAGACCTCCAGCAACCCCGCTGTGGGACCT 600
Db      541 TCAGTCCCCGATCCACAACCTCTCGGAGAGACCTCCAGCAACCCCGCTGTGGGACCT 600

QY      601 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATACGAAGGCGCGCAGGGA 660
|||||
```

```
Db      601 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAAGGCGCGCAGGA 660
|||||
QY      661 GTGGGTAATGCTCAGGAAATTTGGCAATTCGCATTCACATGCTGGGCGACAGAGTCATC 720
|||||
Db      661 GTGGGTAATGCTCAGGAAATTTGGCAATTCGCATTCACATGCTGGGCGACAGAGTCATC 720
|||||
QY      721 ACCACGAGCACCAGCAGCTGGGCTTGGCCACCTACAAATAACACCTCTTCAAGAAATC 780
|||||
Db      721 ACCACGAGCACCAGCAGCTGGGCTTGGCCACCTACAAATAACACCTCTTCAAGAAATC 780
|||||
QY      781 TCCAGTGTCTTCAACGGGGCCAGCAACGACACCACTACTTTCGGCTACGACACCCCTCG 840
|||||
Db      781 TCCAGTGTCTTCAACGGGGCCAGCAACGACACCACTACTTTCGGCTACGACACCCCTCG 840
|||||
QY      841 GGGTATTTGATTTCAACAGATTCACATGCCCACTTTTACACGCTGACTGGCAGCAGCTC 900
|||||
Db      841 GGGTATTTGATTTCAACAGATTCACATGCCCACTTTTACACGCTGACTGGCAGCAGCTC 900
|||||
QY      901 ATCAACAACAATTTGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
|||||
Db      901 ATCAACAACAATTTGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
|||||
QY      961 GTCAAGGAGGTCAAGCAGATGATGGCGTCAAAACCATCGCTTAATTAACCTTACCAGCAG 1020
|||||
Db      961 GTCAAGGAGGTCAAGCAGATGATGGCGTCAAAACCATCGCTTAATTAACCTTACCAGCAG 1020
|||||
QY      1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTTCCGTCTCTGGCTCTGGCACCAG 1080
|||||
Db      1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTTCCGTCTCTGGCTCTGGCACCAG 1080
|||||
QY      1081 GGTGCTCTCTCTCGGCTTCCGGCGGAGCTGTTTTCATGATTCGGCAATACGCTACCTGACG 1140
|||||
Db      1081 GGTGCTCTCTCTCGGCTTCCGGCGGAGCTGTTTTCATGATTCGGCAATACGCTACCTGACG 1140
|||||
QY      1141 CTCACAATTTGGCAGCCGAGCTGTTTCATGATTCGCAATACGCTACCTGCAATATTTCCCT 1200
|||||
Db      1141 CTCACAATTTGGCAGCCGAGCTGTTTCATGATTCGCAATACGCTACCTGCAATATTTCCCT 1200
|||||
QY      1201 TCTCAGATCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
|||||
Db      1201 TCTCAGATCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
|||||
QY      1261 TTCCACAGCAGCTACCGGCACAGCAGCTGGACCGGCTGATGAATCTCTCATCGAC 1320
|||||
Db      1261 TTCCACAGCAGCTACCGGCACAGCAGCTGGACCGGCTGATGAATCTCTCATCGAC 1320
|||||
QY      1321 CAATACCTGTATTAACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGGAC 1380
|||||
Db      1321 CAATACCTGTATTAACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGGAC 1380
|||||
QY      1381 TTGCTGTTTTCAGCGTGGGTCTCCAGCTGGCATGCTTTCAGCCCAAAACTGGCTACCT 1440
|||||
Db      1381 TTGCTGTTTTCAGCGTGGGTCTCCAGCTGGCATGCTTTCAGCCCAAAACTGGCTACCT 1440
|||||
QY      1441 GGACCTCTGTTATCGGCGAGCGCGTTCCTTAAACCAAAACAGCAACAAACAGCAAT 1500
|||||
Db      1441 GGACCTCTGTTATCGGCGAGCGCGTTCCTTAAACCAAAACAGCAACAAACAGCAAT 1500
|||||
QY      1501 TTTTACCTGAGTGGTCTTCAAAATATTAACCTCAATGGGGTGAATTCATCATCAACCT 1560
|||||
Db      1501 TTTTACCTGAGTGGTCTTCAAAATATTAACCTCAATGGGGTGAATTCATCATCAACCT 1560
|||||
QY      1561 GGCACCTGCTATGGCTCTCACAAAGACGAGAGCAAGTTCCTTCCATGACGGGTGC 1620
|||||
Db      1561 GGCACCTGCTATGGCTCTCACAAAGACGAGAGCAAGTTCCTTCCATGACGGGTGC 1620
|||||
QY      1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCTATGGCAATGTGATGAT 1680
|||||
Db      1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCTATGGCAATGTGATGAT 1680
|||||
QY      1681 ACAGACGAAGAGAAATTAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGGACCGGTG 1740
|||||
```

Db	1681	ACAGACGAGAGGAAATTAAGCCACTAAACCTGTGGCCACCGAAGATTTCGGACCGTG	1740	Db	121	GACGGCGGGGTCTGGTCTTCTGGCTTACAACTCGGACCTTCGACACCTTCAACGGACTCGAC	180
Qy	1741	GCAGTCAATTTCCAGAGCAGCAGCAGACCTCTGCGACCGGAGATGTGATCTATGGGA	1800	Qy	181	AAGGGGAGGCCCCGTCAACGCGCGGAGACGACGGCCCTTCGAGACGACAAAGGCTTACGAC	240
Db	1741	GCAGTCAATTTCCAGAGCAGCAGCAGACCTCTGCGACCGGAGATGTGATCTATGGGA	1800	Db	181	AAGGGGAGGCCCCGTCAACGCGCGGAGACGACGGCCCTTCGAGACGACAAAGGCTTACGAC	240
Qy	1801	GCATTTACCTGGCATGGTGGCAAGATAGAGACGTGTACTCTGACGGTCCCAATTTGGGCC	1860	Qy	241	CAGCAGCTCAAGCGGGGTGACAACTCCGTACCTGCGGTATTAACCAACGCGCGAGTTC	300
Db	1801	GCATTTACCTGGCATGGTGGCAAGATAGAGACGTGTACTCTGACGGTCCCAATTTGGGCC	1860	Db	241	CAGCAGCTCAAGCGGGGTGACAACTCCGTACCTGCGGTATTAACCAACGCGCGAGTTC	300
Qy	1861	AAAAATTCCTCACACAGATGGACACTTTTCAACCCGTCTCTCTTATGGGGGCTTTGGACTC	1920	Qy	301	CAGAGCGCTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Db	1861	AAAAATTCCTCACACAGATGGACACTTTTCAACCCGTCTCTCTTATGGGGGCTTTGGACTC	1920	Db	301	CAGAGCGCTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Qy	1921	AAGAACCCGCTCTCAGATCTCATCAAAAAACAGCCGTCTTCTGCGAATCCTCCGGCG	1980	Qy	361	GCCAAGAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGGCGCTTAAAGCAGGCTCT	420
Db	1921	AAGAACCCGCTCTCAGATCTCATCAAAAAACAGCCGTCTTCTGCGAATCCTCCGGCG	1980	Db	361	GCCAAGAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGGCGCTTAAAGCAGGCTCT	420
Qy	1981	GAGTTTCAGCTACAAAGTTTCTTCTTATTCATCATCAACCCATCTCCACAGCAAGTGAGT	2040	Qy	421	GGGAAGAAACGTCGCGGTAGACAGTCGCGCAACAGAGCCAGACTCTCTCTCGGGCATCGGC	480
Db	1981	GAGTTTCAGCTACAAAGTTTCTTCTTATTCATCATCAACCCATCTCCACAGCAAGTGAGT	2040	Db	421	GGGAAGAAACGTCGCGGTAGACAGTCGCGCAACAGAGCCAGACTCTCTCTCGGGCATCGGC	480
Qy	2041	GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCGAG	2100	Qy	481	AAGCAGGCCAGCAGCCGCTTAAAAAGAGACTCAATTTTGGTTCAGACTTGGCGACTCAGAG	540
Db	2041	GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCGAG	2100	Db	481	AAGCAGGCCAGCAGCCGCTTAAAAAGAGACTCAATTTTGGTTCAGACTTGGCGACTCAGAG	540
Qy	2101	TACACATCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTT	2160	Qy	541	TCAGTCCCGCATCCAACTCTCGAGAACTCTCCAGCAACCCCGCTGCTGTGGGACCT	600
Db	2101	TACACATCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTT	2160	Db	541	TCAGTCCCGCATCCAACTCTCGAGAACTCTCCAGCAACCCCGCTGCTGTGGGACCT	600
Qy	2161	TATACAGCCCTGCCCCATTTGGCACCCTGTTACCTTACCCGTCCTCTGT	2209	Qy	601	ACTCAATTTGGCTCAGGCGGTGGCGCAACCAATGGCAGACCAATAGCAAGGCGCCGACCGA	660
Db	2161	TATACAGCCCTGCCCCATTTGGCACCCTGTTACCTTACCCGTCCTCTGT	2209	Db	601	ACTCAATTTGGCTCAGGCGGTGGCGCAACCAATGGCAGACCAATAGCAAGGCGCCGACCGA	660
RESULT 3				Qy	661	GTGGTAATGCTCAGGAAATTTGGCATTTGCCATTTCCATGCTGGGCGGACAGAGTCAATC	720
LOCUS	CS073614	2211 bp	DNA	linear	PAT 05-MAY-2005		
DEFINITION	Sequence 202 from Patent WO200503321.			Db	661	GTGGTAATGCTCAGGAAATTTGGCATTTGCCATTTCCATGCTGGGCGGACAGAGTCAATC	720
ACCESSION	CS073614			Qy	721	ACCACGACCCGCGACCTGGGCGCTTGGCCACCTTACAATAACCACTCTCAAGCAAAATC	780
VERSION	CS073614.1	GI:63090494		Db	721	ACCACGACCCGCGACCTGGGCGCTTGGCCACCTTACAATAACCACTCTCAAGCAAAATC	780
KEYWORDS	unidentified			Qy	781	TCCAGTGTCTCAACGGGGCCAGCAACGACCACTTCTCGGCTACAGCACCCCTGG	840
SOURCE	unidentified			Db	781	TCCAGTGTCTCAACGGGGCCAGCAACGACCACTTCTCGGCTACAGCACCCCTGG	840
ORGANISM	unclassified.			Qy	841	GGGTATTTTGATTTTCAACAGATTTCACCTGGCACTTTTCAACACGTCGTCGAGCGACTC	900
REFERENCE	1	Wilson, J.M., Gao, G., Alvira, M.R. and Vandenberghe, L.H.		Db	841	GGGTATTTTGATTTTCAACAGATTTCACCTGGCACTTTTCAACACGTCGTCGAGCGACTC	900
AUTHORS		Adeno-associated virus (aav) clades, sequences, vectors containing		Qy	901	ATCAACAAATTTGGGATTCGCGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAA	960
TITLE		same, and uses therefor		Db	901	ATCAACAAATTTGGGATTCGCGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAA	960
JOURNAL		Patent: WO 200503321-A 202 14-APR-2005;		Qy	961	GTCAAGGAGTCAACGCAATGTATGGCGTCAACCACTCGCTTAATACTTACAGCAGC	1020
		The Trustees of the University of Pennsylvania (US)		Db	961	GTCAAGGAGTCAACGCAATGTATGGCGTCAACCACTCGCTTAATACTTACAGCAGC	1020
FEATURES		Location/Qualifiers		Qy	1021	GTTCAGTCTTCTCGGACTCGGAGTACGAGTTCGCTTCCGTCCTCGGCTCTGCGCACCCAG	1080
source		1. .2211		Db	1021	GTTCAGTCTTCTCGGACTCGGAGTACGAGTTCGCTTCCGTCCTCGGCTCTGCGCACCCAG	1080
		/organism="unidentified"		Qy	1081	GGTCCCTCCCTCGTTCGCGCGGACGTTGTCATGATTCGCAATACGGCTACCTGAGC	1140
		/mol_type="unassigned DNA"		Db	1081	GGTCCCTCCCTCGTTCGCGCGGACGTTGTCATGATTCGCAATACGGCTACCTGAGC	1140
		/db_xref="taxon:32644"		Qy	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTCTTCTTCTGCTCGGAAATTTTCCCT	1200
		/note="adeno-associated virus, serotype 1"		Db	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTCTTCTTCTGCTCGGAAATTTTCCCT	1200
ORIGIN				Qy	1201	TCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT	1260
Query Match	100.0%	Score 2209;	DB 6;	Length 2211;			
Best Local Similarity	100.0%	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Matches 2209;	Conservative 0;						
Qy	1	ATGGCTGCCGATGGTTATCTTCAGATTTGGCTCGAGGACAACTCTCTGAGGGCATTGCG	60	Qy	1081	GGTCCCTCCCTCGTTCGCGCGGACGTTGTCATGATTCGCAATACGGCTACCTGAGC	1140
Db	1	ATGGCTGCCGATGGTTATCTTCAGATTTGGCTCGAGGACAACTCTCTGAGGGCATTGCG	60	Db	1081	GGTCCCTCCCTCGTTCGCGCGGACGTTGTCATGATTCGCAATACGGCTACCTGAGC	1140
Qy	61	GAGTGTGGGACTTGAAACCTGGAGCCCGAAGCCCAAGCCCAACAGCAAAAGCAGGAC	120	Qy	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTCTTCTTCTGCTCGGAAATTTTCCCT	1200
Db	61	GAGTGTGGGACTTGAAACCTGGAGCCCGAAGCCCAAGCCCAACAGCAAAAGCAGGAC	120	Db	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTCTTCTTCTGCTCGGAAATTTTCCCT	1200
Qy	121	GACGGCGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC	180	Qy	1201	TCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT	1260
Db	121	GACGGCGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC	180	Db	1201	TCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT	1260

QY	1261	TTCCACAGCAGCTACGCGACAGCCAGAGCCTGGACCGCGTGATGAATCCTCTCATCGAC	1320
Db	1261		
QY	1321	CAATACCTGTATACCTGACAGAACTCAAAATCAGTCGGAGTCCCAAAACAGGAC	1380
Db	1321		
QY	1381	TTGCTGTTTAGCGTGGGTCTCCAGCTGGCATGCTGTTCAGGCCAAAACCTGGCTACCT	1440
Db	1381		
QY	1441	GGACCTGTATTCGGCAGCAGCGCGTTTCTAAACAAACACAGACAAACAGCAAT	1500
Db	1441		
QY	1501	TTTACCTGAGCTGGTCTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAACCT	1560
Db	1501		
QY	1561	GGCAGCTGTATGSCCTCACAAAGACGACGAGCAAGTTCTTTCCCATGAGCGGTGC	1620
Db	1561		
QY	1621	ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGACAAATGTATGATT	1680
Db	1621		
QY	1681	ACAGACGAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTG	1740
Db	1681		
QY	1741	CGAGTCAATTCAGAGCAGCAGCAGACCCCTGCGACCGAGATGTGATGCTATGGGA	1800
Db	1741		
QY	1801	GCAATTACCTGGCATGGTGGCAAGATAGAGACGTGTACCTGCAGGTCCTCATTTGGGC	1860
Db	1801		
QY	1861	AAATTCCTCACACAGATGACACTTTACCCGTCTCTTTATGGGGGCTTTGGAATC	1920
Db	1861		
QY	1921	AGAAACCGCCTCCTCAGATCCTCATCAAAACACCGCTGTCTCGAATCCTCCGGCG	1980
Db	1921		
QY	1981	GAGTTTTCAGCTACAAAGTTTGCTTCATTTCATCAACCAATACCTCCACAGGACAAGTGAGT	2040
Db	1981		
QY	2041	GTGGAATTAATGGAGTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCGGAATGCGAG	2100
Db	2041		
QY	2101	TACACATCAATTAATGCAAAATCTGCCAAGTTGATTTTACTTGGACAAATGCACTT	2160
Db	2101		
QY	2161	TATACTGAGCCTCGCCCCATTTGGCACCGCTTACCTTACCCGTCCTCCCTGT	2209
Db	2161		

RESULT 4
AR562503 LOCUS 2211 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 12 from patent US 6759237.
ACCESSION AR562503
VERSION AR562503.1 GI:53976569
KEYWORDS

SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 2211)
TITLE	Wilson, J.M. and Xiao, W.
JOURNAL	Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same
FEATURES	Patent: US 6759237-A 12 06-JUL-2004;
source	The Trustees of the University of Pennsylvania; Philadelphia, PA
	Location/Qualifiers
	1..2211
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Query Match	100.0%; Score 2209; DB 6; Length 2211;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2209; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGGCTGCCGATGGTTATCTTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60
Db	1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60
Qy	61 GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCAAACAGCAAGCAGGAC 120
Db	61 GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCCAACAGCAAGCAGGAC 120
Qy	121 GACGGCCGGGTCTGGTGTCTTCTGGCTACAAAGTACCTCGGACCCCTTCAACGGACTCGAC 180
Db	121 GACGGCCGGGTCTGGTGTCTTCTGGCTACAAAGTACCTCGGACCCCTTCAACGGACTCGAC 180
Qy	181 AAGGGGAGCCCGTCAACGCGGGGAGCGAGCGGCGCTCGAGCAGCAAGGCTTACGAC 240
Db	181 AAGGGGAGCCCGTCAACGCGGGGAGCGAGCGGCGCTCGAGCAGCAAGGCTTACGAC 240
Qy	241 CAGCAGCTCAAGCGGGTGACAACTCCGTACTCGGTATACCAACGCGCGCGAGTTT 300
Db	241 CAGCAGCTCAAGCGGGTGACAACTCCGTACTCGGTATACCAACGCGCGCGAGTTT 300
Qy	301 CAGGAGCGTCTGCAAGAAGATACGCTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCAG 360
Db	301 CAGGAGCGTCTGCAAGAAGATACGCTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCAG 360
Qy	361 GCCAAGAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCT 420
Db	361 GCCAAGAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCT 420
Qy	421 GGAAGAAGCGTCCGGTAGAGCAGTCGCCAAGAGCCAGACTCCTCTCGGGCATCGGC 480
Db	421 GGAAGAAGCGTCCGGTAGAGCAGTCGCCAAGAGCCAGACTCCTCTCGGGCATCGGC 480
Qy	481 AAGACAGGCGCAGCGCCGTAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540
Db	481 AAGACAGGCGCAGCGCCGTAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540
Qy	541 TCAGTCCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGTGTGGGACCT 600
Db	541 TCAGTCCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGTGTGGGACCT 600
Qy	601 ACTCAATGGCTTCAGCGGTGGCGCAACATGGCAGACAAATACGAAGCGCGCAGCGGA 660
Db	601 ACTCAATGGCTTCAGCGGTGGCGCAACATGGCAGACAAATACGAAGCGCGCAGCGGA 660
Qy	661 GTGGGTAATGCCTCAGGAAATGGCAATTCGATTTCCCATGGCTGGGCGACAGAGTCAATC 720
Db	661 GTGGGTAATGCCTCAGGAAATGGCAATTCGATTTCCCATGGCTGGGCGACAGAGTCAATC 720
Qy	721 ACCACAGCAGCCCGCAGCTTGGGCTTGGCCACTTACCAATACCACTCTTACAAGCAAAATC 780
Db	721 ACCACAGCAGCCCGCAGCTTGGGCTTGGCCACTTACCAATACCACTCTTACAAGCAAAATC 780
Qy	781 TCCAGTGTCTTCAACGGGGGCGAGCAACGAACTACTTTCGGGTACAGCAGCCCTTGG 840

781 TCCAGTGTCTTCAACGGGGCCAGCAACGACCAACCACTACTTCGGCTACAGCACCCCTGG 840
841 GGGTATTTTGAATTTCAACAGATTTCACTGCCACTTTTTCACCAAGTGAATGGCAGGACATC 900
841 GGGTATTTTGAATTTCAACAGATTTCACTGCCACTTTTTCACCAAGTGAATGGCAGGACATC 900
901 ATCAACAACAATTTGGGATTCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
901 ATCAACAACAATTTGGGATTCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
961 GTCAAGAGAGTCAAGAGCAATGATGGCGTCAACCAATCGCTTAATAACCTTACCAAGCAG 1020
961 GTCAAGAGAGTCAAGAGCAATGATGGCGTCAACCAATCGCTTAATAACCTTACCAAGCAG 1020
1021 GTTCAAGTCTTTCGAGCTCGGAGTACCAAGCTTTCGCTACGTCTCGGCTCTCGCACCA 1080
1021 GTTCAAGTCTTTCGAGCTCGGAGTACCAAGCTTTCGCTACGTCTCGGCTCTCGCACCA 1080
1081 GGCTGCTCCCTCCGTTCCCGGGGAGCGTTTCATGATTCGGCAATACGGCTACCTGAGC 1140
1081 GGCTGCTCCCTCCGTTCCCGGGGAGCGTTTCATGATTCGGCAATACGGCTACCTGAGC 1140
1141 CTCAACAATGGCAGCAAGCGTGGAGCGTTTCATCTCTTTTACTGCTGGAATATTTCCCT 1200
1141 CTCAACAATGGCAGCAAGCGTGGAGCGTTTCATCTCTTTTACTGCTGGAATATTTCCCT 1200
1201 TCTCAGATGCTGAGAACGGGCAACAATTTTACCTTCAGCTACACCTTTGAGGAAGTGCT 1260
1201 TCTCAGATGCTGAGAACGGGCAACAATTTTACCTTCAGCTACACCTTTGAGGAAGTGCT 1260
1261 TTCCACAGAGCTACGCGCACAGCCAGAGCTGGACCGCTGATGATTCCTCTCATCGAC 1320
1261 TTCCACAGAGCTACGCGCACAGCCAGAGCTGGACCGCTGATGATTCCTCTCATCGAC 1320
1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGGAC 1380
1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGGAC 1380
1381 TTGCTGTTTGAAGCGTGGTCTCCAGCTGCGATGTCTGTTTCAAGCCCAAAACTGGCTACCT 1440
1381 TTGCTGTTTGAAGCGTGGTCTCCAGCTGCGATGTCTGTTTCAAGCCCAAAACTGGCTACCT 1440
1441 GGAACCTGTATTTCGGCAGCAGCGGTTCTTAAACAAACAGACAGCAACCAACAGCAAT 1500
1441 GGAACCTGTATTTCGGCAGCAGCGGTTCTTAAACAAACAGACAGCAACCAACAGCAAT 1500
1501 TTTACCTGAGCTGGTGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCT 1560
1501 TTTACCTGAGCTGGTGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCT 1560
1561 GGCATGCTATGGCTTCAACAAGACGACGAAAGCAAGTTCCTTCCCATGAGCGGTGTC 1620
1561 GGCATGCTATGGCTTCAACAAGACGACGAAAGCAAGTTCCTTCCCATGAGCGGTGTC 1620
1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCTGCAATGGACAATGTCATGATT 1680
1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCTGCAATGGACAATGTCATGATT 1680
1681 ACAGACGAAGAGAAATTAAGACCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
1681 ACAGACGAAGAGAAATTAAGACCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
1741 GCAGTCAATTTTCAGAGCAGCAGACAGACCTTGCACCGGAGATGTGATGCTATGGGA 1800
1741 GCAGTCAATTTTCAGAGCAGCAGACAGACCTTGCACCGGAGATGTGATGCTATGGGA 1800
1801 GCATTACCTGGCATGTGTGGCAAGATAGAGAGTGTACTGACAGGTTCCATTTGGGCC 1860
1801 GCATTACCTGGCATGTGTGGCAAGATAGAGAGTGTACTGACAGGTTCCATTTGGGCC 1860
1861 AAAATTTCTCACACAGATGACACTTTTCAACCGGTCTCTTATGGCGGCTTTGGACTC 1920
1861 AAAATTTCTCACACAGATGACACTTTTCAACCGGTCTCTTATGGCGGCTTTGGACTC 1920

QY 1921 AAGAACCCGCTCTCAGATCCTCATCAAAAACAGCGCTGTTCCTGCGAATCCTCCGGCG 1980
Db 1921 AAGAACCCGCTCTCAGATCCTCATCAAAAACAGCGCTGTTCCTGCGAATCCTCCGGCG 1980
QY 1981 GAGTTTTTCAGCTACAAAGTTTGTCTTCAATTCACCCCAATCTCCACAGGACAAGTGAGT 2040
Db 1981 GAGTTTTTCAGCTACAAAGTTTGTCTTCAATTCACCCCAATCTCCACAGGACAAGTGAGT 2040
QY 2041 GTGGAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGTGGAAATCCCGAAGTGAG 2100
Db 2041 GTGGAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGTGGAAATCCCGAAGTGAG 2100
QY 2101 TACATATCAATATGCAAAATCTGCCAAAGTTTGTGATTTTACTGTGGACAACATGGACTT 2160
Db 2101 TACATATCAATATGCAAAATCTGCCAAAGTTTGTGATTTTACTGTGGACAACATGGACTT 2160
QY 2161 TATACGTAGCTCGCCCAATTTGGACACCGTTTACCTTACCCGTCCCTGT 2209
Db 2161 TATACGTAGCTCGCCCAATTTGGACACCGTTTACCTTACCCGTCCCTGT 2209

RESULT 5
BD242766
LOCUS BD242766
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector and host cell containing the same.
ACCESSION BD242766
VERSION BD242766.1 GI:33052536
KEYWORDS JP 2002529098-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson, J.M. and Xiao, W.
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector and host cell containing the same
JOURNAL Patent: JP 2002529098-A 1 10-SEP-2002;
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT OS AAV-1
PN JP 2002529098-A/1
PD 10-SEP-2002
PF 02-NOV-1999 JP 2000581227
PR 05-NOV-1998 US 60/107114
PI JAMES M WILSON, WEIDONG XIAO
PC C12N15/09, A61K31/711, A61K48/00, A61P43/00, C12N1/15, C12N1/19, PC C12N1/21,
PC C12N5/10, C12P21/02//A61K35/76, C12N15/00, C12N5/00 CC
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH Key Location/Qualifiers
FT CDS (335)..(2206)
FT CDS (2223)..(4430).
FEATURES Location/Qualifiers
source 1..4718
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 2209; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 60
Db 2223 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 2282
QY 61 GAGTGTGGGACTTGAACCTTGAGACCCCGAGCCCAAGCCCAAGCCCAAGCCCAAGCAGGAC 120
Db 2283 GAGTGTGGGACTTGAACCTTGAGACCCCGAGCCCAAGCCCAAGCCCAAGCAGGAC 2342

QY 121 GACGGCCGGGTCTGGTCTTCTGGCTACAGTACTCTGGACCCCTTCAACGACTCGAC 180
DB 2343 GACGGCCGGGTCTGGTCTTCTGGCTACAGTACTCTGGACCCCTTCAACGACTCGAC 2402
QY 181 AAGGGGAGCCGTCAACCGGGCGGACGAGCGGCCCTCGAGCA CGAAGAAGGCTACGAC 240
DB 2403 AAGGGGAGCCGTCAACCGGGCGGACGAGCGGCCCTCGAGCA CGAAGAAGGCTACGAC 2462
QY 241 CAGCAGCTCAAAAGCGGGTCGAATCCGTACTCTGGGTATAAACCAGCGCGACGCCGAGTTT 300
DB 2463 CAGCAGCTCAAAAGCGGGTGACAATCCGTACTCTGGGTATAAACCAGCGCGACGCCGAGTTT 2522
QY 301 CAGGAGCGCTCGCAAGAAGTACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
DB 2523 CAGGAGCGCTCGCAAGAAGTACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2582
QY 361 GCCAAGAAGCGGGTTCTCAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT 420
DB 2583 GCCAAGAAGCGGGTTCTCAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT 2642
QY 421 GGAAGAAGACGTCGGGTAGAGCAGTCGCCACAAGACCGAGACTCTCTCCCTCGGCATCGGC 480
DB 2643 GGAAGAAGACGTCGGGTAGAGCAGTCGCCACAAGACCGAGACTCTCTCCCTCGGCATCGGC 2702
QY 481 AAGACAGGCGCAGAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTCGCGACTCAGAG 540
DB 2703 AAGACAGGCGCAGAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTCGCGACTCAGAG 2762
QY 541 TCAGTCCCAGATCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGACCT 600
DB 2763 TCAGTCCCAGATCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGACCT 2822
QY 601 ACTACAAATGGCTTCAGCGGTGCGCACCAATGGCAGACAATAAACGAAGCGCCGACGGA 660
DB 2823 ACTACAAATGGCTTCAGCGGTGCGCACCAATGGCAGACAATAAACGAAGCGCCGACGGA 2882
QY 661 GTGGGTAAATGCCCTCAGGAATTTGGCATTTGGCATTTCCACATGGCTGGGCGACAGATCATC 720
DB 2883 GTGGGTAAATGCCCTCAGGAATTTGGCATTTGGCATTTCCACATGGCTGGGCGACAGATCATC 2942
QY 721 ACCACAGCAACCCGACCTTGGGCTTGGCCACCTACAATAACCACTCTACAAGCAATC 780
DB 2943 ACCACAGCAACCCGACCTTGGGCTTGGCCACCTACAATAACCACTCTACAAGCAATC 3002
QY 781 TCCAGTGTCTCAACGGGGCCAGCAACCACTACTTCCGCTACAGCACCCCTCG 840
DB 3003 TCCAGTGTCTCAACGGGGCCAGCAACCACTACTTCCGCTACAGCACCCCTCG 3062
QY 841 GGGTATTTTGAATTTCAACAGATTTCCACTGGCACTTTTCCACAGTGACTGGCAGCGACTC 900
DB 3063 GGGTATTTTGAATTTCAACAGATTTCCACTGGCACTTTTCCACAGTGACTGGCAGCGACTC 3122
QY 901 ATCAACAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTTCAACATCCAA 960
DB 3123 ATCAACAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTTCAACATCCAA 3182
QY 961 GTCAAGAGGTACGCAAGATGTGGCTCACAACCACTCGCTAATACCTTTACCGACAG 1020
DB 3183 GTCAAGAGGTACGCAAGATGTGGCTCACAACCACTCGCTAATACCTTTACCGACAG 3242
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAACCTCTCGCTCTGGCTCGGCGACCG 1080
DB 3243 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAACCTCTCGCTCTGGCTCGGCGACCG 3302
QY 1081 GGCTGCTCCTCCGTTCCTGGGGAGCTGTTCATGATTCGGAATACGGCTACCTGACG 1140
DB 3303 GGCTGCTCCTCCGTTCCTGGGGAGCTGTTCATGATTCGGAATACGGCTACCTGACG 3362
QY 1141 CTCACAAATGGCAGCGAAGCGTGGGAGCTTCATCTTTTACTCGCTGGTAATTTCCCT 1200
DB 3363 CTCACAAATGGCAGCGAAGCGTGGGAGCTTCATCTTTTACTCGCTGGTAATTTCCCT 3422
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCT 1260

DB 3423 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCT 3482
QY 1261 TTCCACAGCAGCTACGCGCACAGCCAGAGCCTTGGACCGGCTGATGAATCTCTCATCGAC 1320
DB 3483 TTCCACAGCAGCTACGCGCACAGCCAGAGCCTTGGACCGGCTGATGAATCTCTCATCGAC 3542
QY 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAAAGAC 1380
DB 3543 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAACAAAGAC 3602
QY 1381 TTGCTGTTTACCGGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACCTGGCTACCT 1440
DB 3603 TTGCTGTTTACCGGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACCTGGCTACCT 3662
QY 1441 GGACCTGTATTTCGGCAGCAGCGGCTTTCTAAAAAATAAAACAGAAACAGAAACAGCAAT 1500
DB 3663 GGACCTGTATTTCGGCAGCAGCGGCTTTCTAAAAAATAAAACAGAAACAGAAACAGCAAT 3722
QY 1501 TTTTACCTGGACTGGTCTTCAAAATATAAACCCTCAATGGCGGTGAATTCATCATCAACCT 1560
DB 3723 TTTTACCTGGACTGGTCTTCAAAATATAAACCCTCAATGGCGGTGAATTCATCATCAACCT 3782
QY 1561 GGCACTGCTATGGCTTCAACAAAGACGACGAAAGTTCCTTTCCCATGAGCGGTGC 1620
DB 3783 GGCACTGCTATGGCTTCAACAAAGACGACGAAAGTTCCTTTCCCATGAGCGGTGC 3842
QY 1621 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAAACACTTGCATTTGGACAATGTCTATGAT 1680
DB 3843 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAAACACTTGCATTTGGACAATGTCTATGAT 3902
QY 1681 ACAGACGAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
DB 3903 ACAGACGAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 3962
QY 1741 GAGTCAATTTTCAGAGCAGCAGACAGACCTTGGACCGGAGATGTGCTATGGGA 1800
DB 3963 GCAGTCAATTTTCAGAGCAGCAGACAGACCTTGGACCGGAGATGTGCTATGGGA 4022
QY 1801 GCATTTACCTGGCATGTGTGGCAAGATAGAGACGTGTACTTGCAGGGTCCCATTTCGGGC 1860
DB 4023 GCATTTACCTGGCATGTGTGGCAAGATAGAGACGTGTACTTGCAGGGTCCCATTTCGGGC 4082
QY 1861 AAAATTCCTCACACAGATGGACACTTTTCAACCCGTCTCTCTTTATGGCGGCTTTGACTC 1920
DB 4083 AAAATTCCTCACACAGATGGACACTTTTCAACCCGTCTCTCTTTATGGCGGCTTTGACTC 4142
QY 1921 AAGAACCCGCTCTCAGATCTCATCAAAACACAGCCTGTTCTTCTGGGAATCCTCCGGCG 1980
DB 4143 AAGAACCCGCTCTCAGATCTCATCAAAACACAGCCTGTTCTTCTGGGAATCCTCCGGCG 4202
QY 1981 GAGTTTTCAGCTACAAGTTTTCCTTCAATTCATCACCCTAATCTCCACAGGACAAGTGA 2040
DB 4203 GAGTTTTCAGCTACAAGTTTTCCTTCAATTCATCACCCTAATCTCCACAGGACAAGTGA 4262
QY 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGTGGAAATCCGGAAGTGCAG 2100
DB 4263 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGTGGAAATCCGGAAGTGCAG 4322
QY 2101 TACACATCAATTTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGACACAAATGGAATTT 2160
DB 4323 TACACATCAATTTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGACACAAATGGAATTT 4382
QY 2161 TATACTGAGCCTCGCCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGT 2209
DB 4383 TATACTGAGCCTCGCCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGT 4431

RESULT 6

AR562498

LOCUS

DEFINITION

ACCESSION

AR562498

Sequence 1 from patent US 6759237.

AR562498

4718 bp

DNA

linear

PAT 08-OCT-2004

VERSION	AR562498.1	GI:53976564	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 4718)		
AUTHORS	Wilson, J.M. and Xiao, M.		
TITLE	Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same		
JOURNAL	Patent: US 6759237-A 1 06-JUL-2004;		
FEATURES	The Trustees of the University of Pennsylvania; Philadelphia, PA		
source	Location/Qualifiers		
	1..4718		
	/organism="unknown"		
ORIGIN	/mol_type="genomic DNA"		
Query Match	100.0%;	Score 2209;	DB 6; Length 4718;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2209;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGGCAATCGC	60
Db	2223	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGGCAATCGC	2282
QY	61	GAGTGGTGGACTTGAACTCGAGCCCGAAGCCCAAGCCCAACCCAGCAAAAGCAGGAC	120
Db	2283	GAGTGGTGGGACTTGAACTCGAGCCCGAAGCCCAAGCCCAACCCAGCAAAAGCAGGAC	2342
QY	121	GACGGCCGGGCTCGGTGCTCTCGGTACAACTACCTCGGACCTTCGAGCCCTTCAACGGACTCGAC	180
Db	2343	GACGGCCGGGCTCGGTGCTCTCGGTACAACTACCTCGGACCTTCGAGCCCTTCAACGGACTCGAC	2402
QY	181	AAGGGGAGCCCGTCAACCGCGGACGACGCGGCCCTCGAGCAGCAAGAGCCCTACGAC	240
Db	2403	AAGGGGAGCCCGTCAACCGCGGACGACGCGGCCCTCGAGCAGCAAGAGCCCTACGAC	2462
QY	241	CAGCAGCTCAAGCGGGTGACAACTCGGTACCTCGGTATATAACCGCGCAGCGCGAGTTT	300
Db	2463	CAGCAGCTCAAGCGGGTGACAACTCGGTACCTCGGTATATAACCGCGCAGCGCGAGTTT	2522
QY	301	CAGGAGCGTCTGCAAGAATAGTCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG	360
Db	2523	CAGGAGCGTCTGCAAGAATAGTCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG	2582
QY	361	GCCAGAAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCT	420
Db	2583	GCCAGAAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCT	2642
QY	421	GGAAAGAAACGTCGGTAGAGCAGTGCACAGAGCCAGACTCTCTCGGGCATCGGC	480
Db	2643	GGAAAGAAACGTCGGTAGAGCAGTGCACAGAGCCAGACTCTCTCGGGCATCGGC	2702
QY	481	AAGACAGGCGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAGACTGGCAGCTCAGAG	540
Db	2703	AAGACAGGCGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAGACTGGCAGCTCAGAG	2762
QY	541	TCAGTCCCGBATCCACAACCTCTCGGAGAACTTCCAGCAACCCCGCTCTGTGGGACCT	600
Db	2763	TCAGTCCCGBATCCACAACCTCTCGGAGAACTTCCAGCAACCCCGCTCTGTGGGACCT	2822
QY	601	ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGGAGACAATAAGAAAGCGCCGACGGA	660
Db	2823	ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGGAGACAATAAGAAAGCGCCGACGGA	2882
QY	661	GTGGGTAAATGGCTTCAGGAAATGGCAATTCGATTCACATGGCTGGGCGACAGAGTCATC	720
Db	2883	GTGGGTAAATGGCTTCAGGAAATGGCAATTCGATTCACATGGCTGGGCGACAGAGTCATC	2942
QY	721	ACCACGAGCAGCCGCACTTGGGCTTGGCCCACTCAATAAGCACTCTTACAGCAAAATC	780
Db	2943	ACCACGAGCAGCCGCACTTGGGCTTGGCCCACTCAATAAGCACTCTTACAGCAAAATC	3002
QY	781	TCCAGTGCTTCAACGGGGGCCAGCAACGACCAACCACTACTTTCGGCTACAGCACCCCTGG	840
Db	3003	TCCAGTGCTTCAACGGGGGCCAGCAACGACCAACCACTACTTTCGGCTACAGCACCCCTGG	3062
QY	841	GGGTATTTTGTATTTCAACAGAGATTCACCTGCGACCTTTTTCACCAACGCTGACTGGCAGGACTC	900
Db	3063	GGGTATTTTGTATTTCAACAGAGATTCACCTGCGACCTTTTTCACCAACGCTGACTGGCAGGACTC	3122
QY	901	ATCAACAAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA	960
Db	3123	ATCAACAAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA	3182
QY	961	GTCAGGAGGTTCACGACGAATGATGGGTCAACCACTCGCTAATTAACCTTTACCAAGCAG	1020
Db	3183	GTCAGGAGGTTCACGACGAATGATGGGTCAACCACTCGCTAATTAACCTTTACCAAGCAG	3242
QY	1021	GTTCAAGTCTTCTCGGACTCGGAGTACCAAGCTTCGGTACGTCCTCGGCTCTGCGCACCAG	1080
Db	3243	GTTCAAGTCTTCTCGGACTCGGAGTACCAAGCTTCGGTACGTCCTCGGCTCTGCGCACCAG	3302
QY	1081	GGCTGCTCCCTCGGTTCCCGGGGAGCGTGTTCATGATTCGCGCAATACGGCTACTGACG	1140
Db	3303	GGCTGCTCCCTCGGTTCCCGGGGAGCGTGTTCATGATTCGCGCAATACGGCTACTGACG	3362
QY	1141	CTCAACAATGGCAGCCAGCGTGGAGCGTTTCATCTCTTTTACTGCTCGGAATATTTCCCT	1200
Db	3363	CTCAACAATGGCAGCCAGCGTGGAGCGTTTCATCTCTTTTACTGCTCGGAATATTTCCCT	3422
QY	1201	TCTCAGATGCTTGAGAAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTTGAGGAAGTGCT	1260
Db	3423	TCTCAGATGCTTGAGAAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTTGAGGAAGTGCT	3482
QY	1261	TTCCACAGCAGCTACGCGCAACGAGCAGCGCTGAGCAGCGGCTGATGAATCTCTCATCGAC	1320
Db	3483	TTCCACAGCAGCTACGCGCAACGAGCAGCGCTGAGCAGCGGCTGATGAATCTCTCATCGAC	3542
QY	1321	CAATACCTGTATTACTGCAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGAC	1380
Db	3543	CAATACCTGTATTACTGCAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGAC	3602
QY	1381	TTTGTCTTTAGCGTGGGCTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACCT	1440
Db	3603	TTTGTCTTTAGCGTGGGCTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACCT	3662
QY	1441	GGACCTCTGTTATCGGACAGCGGCTTTCTAAAACAAACAAACAGACAAACAAAGCAAT	1500
Db	3663	GGACCTCTGTTATCGGACAGCGGCTTTCTAAAACAAACAAACAGACAAACAAAGCAAT	3722
QY	1501	TTTACCTCGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT	1560
Db	3723	TTTACCTCGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT	3782
QY	1561	GGCAGTCTATGGCTCTCACAAAGACGACGAAAGCAAGTTCCTTCCCATGAGCGGTGTC	1620
Db	3783	GGCAGTCTATGGCTCTCACAAAGACGACGAAAGCAAGTTCCTTCCCATGAGCGGTGTC	3842
QY	1621	ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAAACATGTGCAATGTCATGATT	1680
Db	3843	ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAAACATGTGCAATGTCATGATT	3902
QY	1681	ACAGACGAGAGGAAATTAAGGCCACTAACCTGTGGCCACCGAAGATTTTGGGACCGTG	1740
Db	3903	ACAGACGAGAGGAAATTAAGGCCACTAACCTGTGGCCACCGAAGATTTTGGGACCGTG	3962
QY	1741	GCAGTCAATTTTCCAGCAGCAGCAGCAGACCTTCGCGACCGGAGATGTGCTATGGGA	1800
Db	3963	GCAGTCAATTTTCCAGCAGCAGCAGCAGACCTTCGCGACCGGAGATGTGCTATGGGA	4022
QY	1801	GCATTAACCTGGCATGGTGTGGCAAGATAGAGAGCTGTACCTGACGGGTCCCATTTGGGCC	1860
Db	4023	GCATTAACCTGGCATGGTGTGGCAAGATAGAGAGCTGTACCTGACGGGTCCCATTTGGGCC	4082
QY	1861	AAATTTCTCTCACACAGATGGACACTTTTCAACCGCTCTCTCTTTATGGGGCGGCTTGGACTC	1920

```
Db 4083 AAAATTCCTCACACAGATGACACTTTTCACCCGCTCTCCTCTTATGCGCGCTTGGACTC 4142
Qy 1921 AAGAACCCGCTCCTCAGATCCTCATCAAAAACACGCTGTTCCTCGAATCCTCGCGG 1980
Db 4143 AAGAACCCGCTCCTCAGATCCTCATCAAAAACACGCTGTTCCTCGAATCCTCGCGG 4202
Qy 1981 GAGTTTTTCAGCTACAAAGTTTGCTTCATTCATCACCCCAATACCTCCACAGACAAGTAGT 2040
Db 4203 GAGTTTTTCAGCTACAAAGTTTGCTTCATTCATCACCCCAATACCTCCACAGACAAGTAGT 4262
Qy 2041 GTGGAAATTCATGGGAGCTGCAGAAAGAAAACAGAACGCTGGAAATCCCGAAGTCGAG 2100
Db 4263 GTGGAAATTCATGGGAGCTGCAGAAAGAAAACAGAACGCTGGAAATCCCGAAGTCGAG 4322
Qy 2101 TACACATCCAAATATGCAAAATCTGCCAAGCTTGATTTTACTGTGGACAAATGACATT 2160
Db 4323 TACACATCCAAATATGCAAAATCTGCCAAGCTTGATTTTACTGTGGACAAATGACATT 4382
Qy 2161 TATACTGAGCTCGCCCCCATTTGGCACCCGTTACCTTACCGTCCCTGT 2209
Db 4383 TATACTGAGCTCGCCCCCATTTGGCACCCGTTACCTTACCGTCCCTGT 4431

RESULT 7
AX753251
LOCUS AX753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
ACCESSION AX753251
VERSION AX753251.1 GI:32166108
KEYWORDS Adeno-associated virus 1
SOURCE Adeno-associated virus 1
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE
AUTHORS Gao, G., Wilson, J. M. and Alvira, M.
TITLE A method of detecting and/or identifying adeno-associated virus
(JAV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
source
1. .4718
/organism="Adeno-associated virus 1"
/mol_type="unassigned DNA"
/db_xref="taxon:85106"

ORIGIN
Query Match 100.0%; Score 2209; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTTCG 60
Db 2223 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTTCG 2282
Qy 61 GAGTGGTGGACTTGAACCTCTGAGCCCCGAAAGCCCAAGCCAAACAGCAAGCAAGCAGGAC 120
Db 2283 GAGTGGTGGACTTGAACCTCTGAGCCCCGAAAGCCCAAGCCAAACAGCAAGCAGGAC 2342
Qy 121 GACGGCCGGGTCTGGTCTTCCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 180
Db 2343 GACGGCCGGGTCTGGTCTTCCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 2402
Qy 181 AAGGGGAGCCCTCAACGGGGGAGCGACGCGCCCTCGAGCAGCAAGGCTTAGCAG 240
Db 2403 AAGGGGAGCCCTCAACGGGGGAGCGACGCGCCCTCGAGCAGCAAGGCTTAGCAG 2462
Qy 241 CAGCAGCTCAAGCGGGTGACAAATCCGTACTCGGTATACCAACACGCGCGCGAGTTT 300
Db 2463 CAGCAGCTCAAGCGGGTGACAAATCCGTACTCGGTATACCAACACGCGCGCGAGTTT 2522
Qy 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTTCGGCGGAGCAGTCTCCAG 360
Db 3603 TTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGGCCCAAAAATGGCTACCT 3662
```

```
Db 2523 CAGGAGCGTCTGCAAGAGATAGTCTTTTGGGGGCAACCTTCGGGCGAGCAGTCTTCCAG 2582
Qy 361 GCCAAGAACGGGTTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGTCTCT 420
Db 2583 GCCAAGAACGGGTTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGTCTCT 2642
Qy 421 GGAAGAAACGTTCGGTAGAGCAGTGCACCAAGAGCCAGACTCTCCTCGGCGATCGGC 480
Db 2643 GGAAGAAACGTTCGGTAGAGCAGTGCACCAAGAGCCAGACTCTCCTCGGCGATCGGC 2702
Qy 481 AAGACAGGCGCAGCAGCCCTGAAAGAGACTCAATTTTGGTTCAGACTGCGGACTCAGAG 540
Db 2703 AAGACAGGCGCAGCAGCCCTGAAAGAGACTCAATTTTGGTTCAGACTGCGGACTCAGAG 2762
Qy 541 TCAGTCCCCTGATCCCAACCTCTCGAGAACTCTCCAGCAACCCCTGCTGTGGGACCT 600
Db 2763 TCAGTCCCCTGATCCCAACCTCTCGAGAACTCTCCAGCAACCCCTGCTGTGGGACCT 2822
Qy 601 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGGAGACAATAACGAAGCGCGGACGGA 660
Db 2823 ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGGAGACAATAACGAAGCGCGGACGGA 2882
Qy 661 GTGGGTAAATGCCCTCAGGAAATTTGGCAATTCGATTTCCACATGGCTGGCGGACAGATCATC 720
Db 2883 GTGGGTAAATGCCCTCAGGAAATTTGGCAATTCGATTTCCACATGGCTGGCGGACAGATCATC 2942
Qy 721 ACCACAGCACCCGCACTGGGCTTTGCCCACTCAACAATAACCACTCTCAACAGCAAAATC 780
Db 2943 ACCACAGCACCCGCACTGGGCTTTGCCCACTCAACAATAACCACTCTCAACAGCAAAATC 3002
Qy 781 TCAGTGTCTTCAACGGGGGCGAGCAACCACTACTTTCGGCTACAGCACCCTCTGG 840
Db 3003 TCAGTGTCTTCAACGGGGGCGAGCAACCACTACTTTCGGCTACAGCACCCTCTGG 3062
Qy 841 GGGTATTTTGAATTTCAACAGATTTCACCTGCCACTTTTTCACACGTGACTGGCAGCGACTC 900
Db 3063 GGGTATTTTGAATTTCAACAGATTTCACCTGCCACTTTTTCACACGTGACTGGCAGCGACTC 3122
Qy 901 ATCAACAAACAATTTGGGGATTTCGGGCGCAAGAGACTCAAACTTTCAACATCAACAA 960
Db 3123 ATCAACAAACAATTTGGGGATTTCGGGCGCAAGAGACTCAAACTTTCAACATCAACAA 3182
Qy 961 GTCAAGAGGTTCAGCAGCAATGATGGCGTCACAAACCATCGCTAATAACCTTACGACAGC 1020
Db 3183 GTCAAGAGGTTCAGCAGCAATGATGGCGTCACAAACCATCGCTAATAACCTTACGACAGC 3242
Qy 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAGTTCCTCGTACGTCTCGGCTCTGCGCACAG 1080
Db 3243 GTTCAAGTCTTCTCGGACTCGGAGTACCAGTTCCTCGTACGTCTCGGCTCTGCGCACAG 3302
Qy 1081 GGCTGCTCTCCTCCGTTCCCGCGGACGTGTTTCATGATTCGCAATACCGCTACCTGACG 1140
Db 3303 GGCTGCTCTCCTCCGTTCCCGCGGACGTGTTTCATGATTCGCAATACCGCTACCTGACG 3362
Qy 1141 CTCAAACAATGGCAGCAAGCGGTGGAGCTTCATTCCTTTTACTGCTCGGTAATTTTCCCT 1200
Db 3363 CTCAAACAATGGCAGCAAGCGGTGGAGCTTCATTCCTTTTACTGCTCGGTAATTTTCCCT 3422
Qy 1201 TCTCAGATGCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
Db 3423 TCTCAGATGCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 3482
Qy 1261 TTCACAGCAGCTACCGCAGCAGCGAGCTTCGAGCGGCTGATGATTCCTCTCATCGAC 1320
Db 3483 TTCACAGCAGCTACCGCAGCAGCGAGCTTCGAGCGGCTGATGATTCCTCTCATCGAC 3542
Qy 1321 CAATACTGTGATTTACTTGAAACAGAACTCAAAAATCAGTCCGGGAAGTGCCCAAAAACAAGGAC 1380
Db 3543 CAATACTGTGATTTACTTGAAACAGAACTCAAAAATCAGTCCGGGAAGTGCCCAAAAACAAGGAC 3602
Qy 1381 TTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGGCCCAAAAATGGCTACCT 1440
Db 3603 TTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGGCCCAAAAATGGCTACCT 3662
```

JOURNAL	Submitted (05-MAY-1998)	IGHG, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA
FEATURES	Location/Qualifiers	
Source	1. 4718	/organism="Adeno-associated virus 1"
		/mol_type="genomic DNA"
		/db_xref="taxon:85106"
		/note="AAV1"
	335. .2206	
CDS	/codon_start=1	
	/product="nonstructural protein"	
	/protein_id="AAD27758.1"	
	/db_xref="GI:4689097"	
	/translation="MPGFYEIVKPSDLDEHLPGISDSFVSVVAARKEMELPPDSMDK LNLTEQAPLTVAEKLDLFVOMRVSQAPEALFFVOFEKESYFHLHILVETTVGVS MVLEGFSLQIDRLVOTIYGLIEPTLNFWEFAVTKTRNGAGGNKVVDCECVINPLLPK TQPELOWANTWMEYISACLNLAERKRLVAQLHTHSOQENKLENPNSDAPVIRAL KTSARYMELVMDVDRGITSBKQWQIDQASYSIFNSAASNSRSQIKAAALDNAGKIMAL TKSPADVLVGAEPADIKTNRIYRILELNGEPAYAGSVFLGWAQKRGKRTNIMLFG PATTKTNIAEIAIAHVPFYGCVMNTNENPFNDCVDKMWIMWEEBGMKTKVSVESKA ILGGSKVRVDQKCKSSAQIDPTPVIYTSNTNMCNAVIDGNSTTTFEHOQPLQDRMFKEFL TRRLHDPKGVTKQEVKEFRMAQDHVEVAHFYVRKGGANKRPADGADKSEPKRA CFSVADPSTSDAEGAPVDPADRYQNKCSRHAGMLQMLFPCKTCTERNQNPINICFTHGT RDCSECFPGVSVSQPVVRKRYRKLCALHLLGRAPEIACADLVNVDLDDCCVSSQ" 2223. .4433	
	/codon_start=1	
	/product="capsid protein"	
	/protein_id="AAD27757.1"	
	/db_xref="GI:4689097"	
	/translation="MAADGYLDPWLENDLSEGIREWMDLKPAPKPAKQKQDDGRG LVLPGYKLYGPNGLDKGEPNVAAALAEHDKAYDQQLKAGDNPYLRVNHADAEPGE RLOEDTSFGNGLRAVFAQAKRVLLEPLGLEVEGATAPGKKRPVPEQSPDSSSGIG KTGQPAKRLNFCGTGDSVPDPQPLGPPPATPAAGVPTTWASGGGA PMADNNRGA DVGNASGNHCDSTLWLDRVITTSRTWALPTYNHLYKQISSASTGASDNHIFYGY STPMGYDFNRFHCFSPRDMQRLLNNWGRFRLNLFNFQIQLVKEVTINDGVTTIA NNLTSYVQFSDSBIYQLPYVLGSAHOGCLPPFADVPFMIPOYGYLTNNGSQVGRSS FYCLTEYFQSMRLRTGNFTFSTYTFEYPFHSSYAHQSQDLNMLPIDQLYLYNLRTS NOSGSAONKDLFLSRGSPAGMSVOPKMLPGPCVQRVRSKTKTDNNNSNFTWTGASK YNLNGRESIINPGTAMASHKDDEDKRPMPGSMVIFGKESAGASNTALDNVMI TDEBEI KATNPATERPGTVAVNFPQSSDTPATGDVHMGALPMVMQDRDVLQOP I WAKIIPH TDGHPHPLMGFGFLGNPPQILIKNTVPFANPPAEFSATKPFASFITQYSTGQSVSE IEWELQKENSXRMNPEVOYTSNVAKSANVDFTVDNNGLYTEPRPIGTLYLTRLP"	
ORIGIN		
Query Match	100.0%;	Score 2209; DB 13; Length 4718;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 2209; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGTGAGGGCATTCCG 60
Db	2223	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGTGAGGGCATTCCG 2282
Qy	61	GAGTGGTGGGACTTGAAACCTTGGAGCCCCGGAAGCCCAAGCCAAACAGCAAAAAGCAGGAC 120
Db	2283	GAGTGGTGGGACTTGAAACCTTGGAGCCCCGGAAGCCCAAGCCAAACAGCAAAAAGCAGGAC 2342
Qy	121	GACGCGCGGGTCTGGTCTCTTCTGGCTTACAAGTACCTTCGGACCCCTTCAACGAGCTCGAC 180
Db	2343	GACGCGCGGGTCTGGTCTCTTCTGGCTTACAAGTACCTTCGGACCCCTTCAACGAGCTCGAC 2402
Qy	181	AAGGGGGAGCCCGTCAACGCGCGGAGCGCAGCGGCCCTTCGAGCACGACAAAGGCCCTACGAC 240
Db	2403	AAGGGGGAGCCCGTCAACGCGCGGAGCGCAGCGGCCCTTCGAGCACGACAAAGGCCCTACGAC 2462
Qy	241	CAGCAGCTCAAAAGCGGGTGACAATCCGTACTCTGGCGTATACCAACGACCGGACGCCGAGTTT 300
Db	2463	CAGCAGCTCAAAAGCGGGTGACAATCCGTACTCTGGCGTATACCAACGACCGGACGCCGAGTTT 2522
Qy	301	CAGGAGCGTCTGCAAGAGATACGCTCTTTTGGGGGCAACCTTCGGGGCAGCAGTCTTCCAG 360
Db	2523	CAGGAGCGTCTGCAAGAGATACGCTCTTTTGGGGGCAACCTTCGGGGCAGCAGTCTTCCAG 2582
Qy	361	GCCAAAGAGCGGGTTCTCGAAACCTCTCGGTCTGGTTGAGGAAGCGCGTCAAGACGGCTCT 420

Db 2583 GGCACAGAGCGGTTCTCGAACCCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 2642
QY 421 GGAAGAAACGTCGGTAGAGAGTCGCCACAGAGCAGACTCCTCTCGGGCATCGGC 480
Db 2643 GGAAGAAACGTCGGTAGAGAGTCGCCACAGAGCAGACTCCTCTCGGGCATCGGC 2702
QY 481 AAGACAGGCGCAGCAGCGCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540
Db 2703 AAGACAGGCGCAGCAGCGCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 2762
QY 541 TCAGTCCCCGATCCAACTCTCGGAGAACTTCAGGAAACCCCGCTGCTGTGGACCT 600
Db 2763 TCAGTCCCCGATCCAACTCTCGGAGAACTTCAGGAAACCCCGCTGCTGTGGACCT 2822
QY 601 ACTACAAATGGCTTCAGCGGTCGCGCACCAATGGCAGACAAATAACGAAGCGCCGACGGA 660
Db 2823 ACTACAAATGGCTTCAGCGGTCGCGCACCAATGGCAGACAAATAACGAAGCGCCGACGGA 2882
QY 661 GTGGGTAATGCTTCAGGAAATGGCAATTCGCAATTCACATGCTGGCGGACAGAGTCATC 720
Db 2883 GTGGGTAATGCTTCAGGAAATGGCAATTCGCAATTCACATGCTGGCGGACAGAGTCATC 2942
QY 721 ACCACAGACCGCGACCTTGGGCTTGGCCACCTACAAATAACCACTCTTACAGCAAAATC 780
Db 2943 ACCACAGACCGCGACCTTGGGCTTGGCCACCTACAAATAACCACTCTTACAGCAAAATC 3002
QY 781 TCCAGTGCTTCAACGGGGCCAGCAACGACCACTACTTGGCTACAGCAACCCCTCG 840
Db 3003 TCCAGTGCTTCAACGGGGCCAGCAACGACCACTACTTGGCTACAGCAACCCCTCG 3062
QY 841 GGGTATTTTGAATTTCAACAGATTCACATGCCACTTTTACACAGCTGACTGGCAGCGACTC 900
Db 3063 GGGTATTTTGAATTTCAACAGATTTCACTGCCACTTTTACCAGCTGACTGGCAGCGACTC 3122
QY 901 ATCAACAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
Db 3123 ATCAACAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 3182
QY 961 GTCAGAGGTCACGAGATGATGGGTCACAAACCATCGCTAATACCTTACCGACAG 1020
Db 3183 GTCAGAGGTCACGAGATGATGGGTCACAAACCATCGCTAATACCTTACCGACAG 3242
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCTCGGCTCTGGGACCCAG 1080
Db 3243 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCTCGGCTCTGGGACCCAG 3302
QY 1081 GGTGCTCTCCCTCGGTCGCGGAGCGTGTTCATGATTTCCGCAATACGCTACCTGACG 1140
Db 3303 GGTGCTCTCCCTCGGTCGCGGAGCGTGTTCATGATTTCCGCAATACGCTACCTGACG 3362
QY 1141 CTCAACAATGGCAGCAGCGGTCGGAGCTTCACTCTTTTACTGCTGGATATTTCCCT 1200
Db 3363 CTCAACAATGGCAGCAGCGGTCGGAGCTTCACTCTTTTACTGCTGGATATTTCCCT 3422
QY 1201 TCTCAGATGCTGAGACGGGCAACAATTTTACCTTCAGCTACACTTTGAGGAGTGCCT 1260
Db 3423 TCTCAGATGCTGAGACGGGCAACAATTTTACCTTCAGCTACACTTTGAGGAGTGCCT 3482
QY 1261 TTCCACAGCAGCTACCGCACAGCCGAGCTTGACCGGCTGATGATTCCTCTCATCGAC 1320
Db 3483 TTCCACAGCAGCTACCGCACAGCCGAGCTTGACCGGCTGATGATTCCTCTCATCGAC 3542
QY 1321 CAATACTGTATTTACTGTAAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGAC 1380
Db 3543 CAATACTGTATTTACTGTAAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGAC 3602
QY 1381 TTGCTGTTTACCGCTGCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACTGGCTACCT 1440
Db 3603 TTGCTGTTTACCGCTGCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACTGGCTACCT 3662
QY 1441 GGAACCTGTATTCCGGCAGCAGCGGTTTCTTAAACAAAAACAGACAAACAGCAAT 1500

Db 3663 GGACCTCTTTATTCGCGACGAGCGCTTTCTTAAAACAAACAAACAGACAAACAGCAAT 3722
QY 1501 TTTTACTCGGACTGGTGTCTTTCAAAATATAACCTCAATGGGCGTGAATTCATCATCAACCT 1560
Db 3723 TTTTACTCGGACTGGTGTCTTTCAAAATATAACCTCAATGGGCGTGAATTCATCATCAACCT 3782
QY 1561 GGCACCTGTATGGCTCTCACAAAAGCAGCAGAAAGTTCCTTTCCCATGAGCGGTGC 1620
Db 3783 GGCACCTGTATGGCTCTCACAAAAGCAGCAGAAAGTTCCTTTCCCATGAGCGGTGC 3842
QY 1621 ATGATTTTGGAAABAGAGAGCGCGGAGCTTCAAAACACTGCAATTGACAAATGTCATGAT 1680
Db 3843 ATGATTTTGGAAABAGAGAGCGCGGAGCTTCAAAACACTGCAATTGACAAATGTCATGAT 3902
QY 1681 ACAGACGAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
Db 3903 ACAGACGAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 3962
QY 1741 GAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCTATGGGA 1800
Db 3963 GCGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCTATGGGA 4022
QY 1801 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGCAGGGTCCCATTTGGGCC 1860
Db 4023 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGCAGGGTCCCATTTGGGCC 4082
QY 1861 AAAATTTCTCACACAGATGGACACTTTTACCCGCTCTCTCTTATGGGCGGCTTTGACTC 1920
Db 4083 AAAATTTCTCACACAGATGGACACTTTTACCCGCTCTCTCTTATGGGCGGCTTTGACTC 4142
QY 1921 AAGAACCCGCTCTCAGATCTCATCAAAAAACGCTGTTCCTGCGAATCTCCTCGGCG 1980
Db 4143 AAGAACCCGCTCTCAGATCTCATCAAAAAACGCTGTTCCTGCGAATCTCCTCGGCG 4202
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCCACAGACAAAGTGA 2040
Db 4203 GAGTTTTCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCCACAGACAAAGTGA 4262
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGGAAGTCAG 2100
Db 4263 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGGAAGTCAG 4322
QY 2101 TACACATCAATTTATGCAAAATCTGCAAAAGTGTGCTGATTTTACTGTGGAACAAATGACTT 2160
Db 4323 TACACATCAATTTATGCAAAATCTGCAAAAGTGTGCTGATTTTACTGTGGAACAAATGACTT 4382
QY 2161 TATACCTGAGCTCGCCCAATTTGGCACCCTTACCTTACCCGTCCTCTGT 2209
Db 4383 TATACCTGAGCTCGCCCAATTTGGCACCCTTACCTTACCCGTCCTCTGT 4431

RESULT 9

AR527492 LOCUS 7447 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 11 from patent US 6723551.
AR527492
ACCESSION AR527492
VERSION AR527492.1 GI:53914590
KEYWORDS

SOURCE

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 7447)

AUTHORS

Kotin, R. M., Urabe, M. and Ding, C.-T.

TITLE

Production of adeno-associated virus in insect cells

JOURNAL

Patent: US 6723551-A 11 20-APR-2004;
The United States of America as represented by the Department of
Health and Human Services; Washington, DC

FEATURES

source
1..7447
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match				99.7%	Score 2202.6	DB 6	Length 7447
Best Local Similarity				99.8%	Pred. No. 0		
Matches 2205				Conservative	0	Mismatches	4
				Indels	0	Gaps	0
Qy	1	ATGGCTGCCGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	60				
Db	4621	ACGGCTGCCGAGCGGTATTCTACCCGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	4680				
Qy	61	GAGTGTGGGACTTGAACCTCGAGCCCCGAAAGCCCAAAGCCCAACACGACAAAGCAGGAC	120				
Db	4681	GAGTGTGGGACTTGAACCTCGAGCCCCGAAAGCCCAAAGCCCAACACGACAAAGCAGGAC	4740				
Qy	121	GACGGCGGGGTCTGGTCTCTTGGCTACAAAGTACCTCGAGCCCTTCAACGGACTCGAC	180				
Db	4741	GACGGCGGGGTCTGGTCTCTTGGCTACAAAGTACCTCGAGCCCTTCAACGGACTCGAC	4800				
Qy	181	RAAGGGGAGCCGTCACCGGGGAGCGAGCGGGCCCTCGAGCAGCAAGGGCTTACGAC	240				
Db	4801	RAAGGGGAGCCGTCACCGGGGAGCGAGCGGGCCCTCGAGCAGCAAGGGCTTACGAC	4860				
Qy	241	CAGCAGCTCAAGCGGGTGACAAATCGTACCTCGGTATTAACACCGCCGACCGCGAGTTT	300				
Db	4861	CAGCAGCTCAAGCGGGTGACAAATCGTACCTCGGTATTAACACCGCCGACCGCGAGTTT	4920				
Qy	301	CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGCAACCTCGGGCAGCAGTCTTCCAG	360				
Db	4921	CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGCAACCTCGGGCAGCAGTCTTCCAG	4980				
Qy	361	GCCAAAGCGGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGGGCGCTTAAGACGGCTCT	420				
Db	4981	GCCAAAGCGGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGGGCGCTTAAGACGGCTCT	5040				
Qy	421	GGAAAGAAAGCTCCGGTAGAGAGTGCACAAAGAGCGAGCTCTCTCGGGGATCGGC	480				
Db	5041	GGAAAGAAAGCTCCGGTAGAGAGTGCACAAAGAGCGAGCTCTCTCTCGGGGATCGGC	5100				
Qy	481	AAGACAGGCGCAGAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGCTGGCGACTCAGAG	540				
Db	5101	AAGACAGGCGCAGAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGCTGGCGACTCAGAG	5160				
Qy	541	TCAGTCCCGATCCACAACTCTCGGAGAACTCCAGCAACCCCCGCTGCTGGGACCT	600				
Db	5161	TCAGTCCCGATCCACAACTCTCGGAGAACTCCAGCAACCCCCGCTGCTGGGACCT	5220				
Qy	601	ACTACAAATGCTTCAGCGGTGCGCACCAATGGCAGACAATAAGAAAGGCGCCGACGGA	660				
Db	5221	ACTACAAATGCTTCAGCGGTGCGCACCAATGGCAGACAATAAGAAAGGCGCCGACGGA	5280				
Qy	661	GTGGGTAAATGCTTCAGGAAATGGCAATTCGATTCACATGGCTGGGCGACAGAGTCAATC	720				
Db	5281	GTGGGTAAATGCTTCAGGAAATGGCAATTCGATTCACATGGCTGGGCGACAGAGTCAATC	5340				
Qy	721	ACCACAGCAACCCGACACTTGGGCTTGGCCCACTCAATTAACCACTCTTCAAGCAAAATC	780				
Db	5341	ACCACAGCAACCCGACACTTGGGCTTGGCCCACTCAATTAACCACTCTTCAAGCAAAATC	5400				
Qy	781	TCCAGTGCTTCAACGGGGGCCAGCAACGACACCACTACTTCCGGCTACAGACCCCCCTGG	840				
Db	5401	TCCAGTGCTTCAACGGGGGCCAGCAACGACACCACTACTTCCGGCTACAGACCCCCCTGG	5460				
Qy	841	GGGTATTTTGATTTCAACAGATTCACATGCCACTTTTACACAGGTGACTGGCAGCGACTC	900				
Db	5461	GGGTATTTTGATTTCAACAGATTCACATGCCACTTTTACACAGGTGACTGGCAGCGACTC	5520				
Qy	901	ATCAACAACAATTTGGGGATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA	960				
Db	5521	ATCAACAACAATTTGGGGATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA	5580				
Qy	961	GTCAAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTTAATACTTACCGACGACG	1020				
Db	5581	GTCAAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTTAATACTTACCGACGACG	5640				
Qy	1021	GTTCAGAGTCTTCTCGGACTCGGAGTACCAAGTTCGGTACGTCCTCGGCTCTCGGCACCCAG	1080				

Db	6721	TACACATCCAAATATATGCAAAATCTGCGCAACGTTGATTTTACGTGTGGAGCAACAATGGACTT	6780
Qy	2161	TATACTAGCCTCGCCCCATTGGCACCCCTTACCTTACCTGCGCTCCCTGT	2209
Db	6781	TATACTAGCCTCGCCCCATTGGCACCCCTTACCTTACCTGCGCTCCCTGT	6829
RESULT 10			
CS073490			
LOCUS	CS073490	2211 bp	DNA linear PAT 05-MAY-2005
DEFINITION	Sequence 78 from Patent WO200503321.		
ACCESSION	CS073490		
VERSION	CS073490.1	GI:63090469	
KEYWORDS	unidentified		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1		
AUTHORS	Wilson,J.M., Gao,G., Alvira,M.R. and Vandenberghe,L.H.		
TITLE	Adeno-associated virus (aav) clades, sequences, vectors containing same, and uses therefor		
JOURNAL	Patent: WO 200503321-A 78 14-APR-2005;		
	The Trustees of the University of Pennsylvania (US)		
FEATURES	Location/Qualifiers		
source	1..2211		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
	/notes="new AAV serotype, clone hu.48"		
ORIGIN			
Query Match	99.4%;	Score 2196.2;	DB 6; Length 2211;
Best Local Similarity	99.6%;	Pred. No. 0;	
Matches 2201;	Conservative 0;	Mismatches 8;	Indels 0; Gaps 0;
Qy	1	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGCATTCGC	60
Db	1	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGCATTCGC	60
Qy	61	GAGTGTGGGACTTGAACTCGAGCCCGGAGCCCAAGGCCAAGCCAGCAAAAGCAGGAC	120
Db	61	GAGTGTGGGACTTGAACTCGAGCCCGGAGCCCAAGGCCAAGCCAGCAAAAGCAGGAC	120
Qy	121	GACGGCGGGCTCTGGTGTCTTCGCTACAAGTACTCGGACCTTCAACGACTCGAC	180
Db	121	GACGGCGGGCTCTGGTGTCTTCGCTACAAGTACTCGGACCTTCAACGACTCGAC	180
Qy	181	AAGGGGAGCCCTCAACCGCGGAGCGAGCGGCCCTCGAGCACGACAAAGGCTACGAC	240
Db	181	AAGGGGAGCCCTCAACCGCGGAGCGAGCGGCCCTCGAGCACGACAAAGGCTACGAC	240
Qy	241	CAGCAGCTCAAGCGGGTGACATCCGTACCTCGGTATATACCAACCGCGCGGAGTTT	300
Db	241	CAGCAGCTCAAGCGGGTGACATCCGTACCTCGGTATATACCAACCGCGCGGAGTTT	300
Qy	301	CAGGAGCGTCTGCAAGAAGTACTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Db	301	CAGGAGCGTCTGCAAGAAGTACTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Qy	361	GCCAAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGGCTAAGACGGTCTCT	420
Db	361	GCCAAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGGCTAAGACGGTCTCT	420
Qy	421	GGAAAGAACTCCGGTAGAGAGTGCACCAAGAGCCAGACTCTCTCGGCGATCGGC	480
Db	421	GGAAAGAACTCCGGTAGAGAGTGCACCAAGAGCCAGACTCTCTCGGCGATCGGC	480
Qy	481	AAGACAGCCAGAGCCCGCTTAAAGAGACTCAATTTTGGTCAGACTGCGGACTCAGAG	540
Db	481	AAGACAGCCAGAGCCCGCTTAAAGAGACTCAATTTTGGTCAGACTGCGGACTCAGAG	540
Qy	541	TCAGTCCCCCATCCAAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGGGACCT	600
Db			

Db	541	TCAGTCCCCCATCCAAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGACCT	600
Qy	601	ACTCAATGGCTTCAGGCGGTGGCGACCAATATGGCAGACAATAACGAAGGCGCCGACGGA	660
Db	601	ACTCAATGGCTTCAGGCGGTGGCGACCAATATGGCAGACAATAACGAAGGCGCCGACGGA	660
Qy	661	GTGGGTAAATGCTCAGGAAATTTGGCATTTCCACATGGCTGGGCGACAGAGTCAATC	720
Db	661	GTGGGTAAATGCTCAGGAAATTTGGCATTTCCACATGGCTGGGCGACAGAGTCAATC	720
Qy	721	ACCACACGACCCGACCTGGGCTTCGCCACCTTACAATAACCACTCTTACAAGCAAAATC	780
Db	721	ACCACACGACCCGACCTGGGCTTCGCCACCTTACAATAACCACTCTTACAAGCAAAATC	780
Qy	781	TCCAGTGTCTTCAACGGGGGCCAGCAACCACTACTTCCGGCTACAGCAACCCCTGG	840
Db	781	TCCAGTGTCTTCAACGGGGGCCAGCAACCACTACTTCCGGCTACAGCAACCCCTGG	840
Qy	841	GGGTATTTTGTATTTCAACAGATTCCACATGGCTTTTACCACGCTGACTGGCAGGACTC	900
Db	841	GGGTATTTTGTATTTCAACAGATTCCACATGGCTTTTACCACGCTGACTGGCAGGACTC	900
Qy	901	ATCAACAACAATTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA	960
Db	901	ATCAACAACAATTTGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA	960
Qy	961	GTCAAGGAGGTCAACGACGAATGATGGCGTCAACCACTCGCTTAATAACCTTACCAGCAG	1020
Db	961	GTCAAGGAGGTCAACGACGAATGATGGCGTCAACCACTCGCTTAATAACCTTACCAGCAG	1020
Qy	1021	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTCTCGGCTCTGGGACACG	1080
Db	1021	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGCTCTCGGCTCTGGGACACG	1080
Qy	1081	GGTCCCTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGGCTACCTGACG	1140
Db	1081	GGTCCCTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGGCTACCTGACG	1140
Qy	1141	CTCAACAATGGCAGCAAGCCGTGGGAGCTTCATCTTTTACTGCTGGAATATTTCCCT	1200
Db	1141	CTCAACAATGGCAGCAAGCCGTGGGAGCTTCATCTTTTACTGCTGGAATATTTCCCT	1200
Qy	1201	TCTCAGATGCTGAGAAAGGGGCAACCTTTTACCTTCAGTACACCTTTGAGGAAGTGCCT	1260
Db	1201	TCTCAGATGCTGAGAAAGGGGCAACCTTTTACCTTCAGTACACCTTTGAGGAAGTGCCT	1260
Qy	1261	TTTCCACAGCAGTACGCGCACAGCAGCGCTGATGAATCCTCTCATCGAC	1320
Db	1261	TTTCCACAGCAGTACGCGCACAGCAGCGCTGATGAATCCTCTCATCGAC	1320
Qy	1321	CAATACCTGTATTAATCTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGAC	1380
Db	1321	CAATACCTGTATTAATCTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAAGGAC	1380
Qy	1381	TTGCTGTTTACGCTGGGTCTCCAGTGGCATGTCTGTTTACGCGCAAACTGGGTACCT	1440
Db	1381	TTGCTGTTTACGCTGGGTCTCCAGTGGCATGTCTGTTTACGCGCAAACTGGGTACCT	1440
Qy	1441	GGACCTGTATTCGGCAGCAGCGCTTCTTAAACAAACAAACAGACAAACAAACAAAT	1500
Db	1441	GGACCTGTATTCGGCAGCAGCGCTTCTTAAACAAACAAACAGACAAACAAACAAAT	1500
Qy	1501	TTTACTCGACTTGGTCTTCAAAATATAACTCAATGGGCGTGAATCCATCATCAACCT	1560
Db	1501	TTTACTCGACTTGGTCTTCAAAATATAACTCAATGGGCGTGAATCCATCATCAACCT	1560
Qy	1561	GGCAGTGTATGGCTTACACAAAGCAGCAAGCAAGTCTTTTCCCATGAGCGGTGTC	1620
Db	1561	GGCAGCGTGTGGCTTACACAAAGCAGCAAGCAAGTCTTTTCCCATGAGCGGTGTC	1620
Qy	1621	ATGATTTTTGAAAAGAGAGCGCGGAGCTTCAACACTGCAATGGCAATGTCATGATT	1680
Db	1621	ATGATTTTTGAAAAGAGAGCGCGGAGCTTCAAGCACTGCAATGGCAATGTCATGATT	1680

QY	1681	ACAGACGAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAGATTTGGGACCGTG	1740
Db	1681	ACAGACGAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAGATTTGGGACCGTG	1740
QY	1741	GCAGTCAATTTCCAGACGACGACGACAGACCTCGACCGGAGATGTGCATCTATGGGA	1800
Db	1741	GCAGTCAATTTCCAGACGACGACGACAGACCTCGACCGGAGATGTGCATCTATGGGA	1800
QY	1801	GCATTACCTGGCATGGTGGCAAGATAGAGACGTGTACTCGAGGTCCTATTTGGGCC	1860
Db	1801	GCATTACCTGGCATGGTGGCAAGATAGAGACGTGTACTCGAGGTCCTATTTGGGCC	1860
QY	1861	AAAAATTCCTACACAGATGACATTTTACCCGCTCTCTTATGGGCGGCTTTGGACTC	1920
Db	1861	AAAAATTCCTACACAGATGACATTTTACCCGCTCTCTTATGGGCGGCTTTGGACTC	1920
QY	1921	AAGAACCCTCCTCAGATCTCATCAAAAACAGCCGTGTTCTTGCAGATCCTCCGGCG	1980
Db	1921	AAGAACCCTCCTCAGATCTCATCAAAAACAGCCGTGTTCTTGCAGATCCTCCGGCG	1980
QY	1981	GAGTTTTCAGCTACAAAGTTTGTCTTCAATTCATCAACCAATCTCCACAGGACAAGTGAGT	2040
Db	1981	GAGTTTTCAGCTACAAAGTTTGTCTTCAATTCATCAACCAATCTCCACAGGACAAGTGAGT	2040
QY	2041	GTGGAAATTAATGGAGCTGCAGAAAGAAAGAAAGCAAGCGCTGGAAATCCGAAATGCAG	2100
Db	2041	GTGGAAATTAATGGAGCTGCAGAAAGAAAGAAAGCAAGCGCTGGAAATCCGAAATGCAG	2100
QY	2101	TACACATCAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAATGGACTT	2160
Db	2101	TACACATCAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAATGGACTT	2160
QY	2161	TATACTGAGCCTGCCCCATTTGGCAACCCGTTACCTTACCCGTCCTCCCTGT	2209
Db	2161	TATACTGAGCCTGCCCCATTTGGCAACCCGTTACCTTACCCGTCCTCCCTGT	2209
RESULT 11			
AY530611			
LOCUS			
DEFINITION	Adeno-associated virus isolate hu.48 capsid protein Vp1 (cap) gene, complete cds.		
ACCESSION	AY530611		
VERSION	AY530611.1 GI:46487868		
KEYWORDS			
SOURCE	Adeno-associated virus		
ORGANISM	Adeno-associated virus		
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.		
AUTHORS	1 (bases 1 to 2211)		
TITLE	Gao,G.; Vandenbergh,L.H., Alvira,M.R., Lu,Y., Calcedo,R., Zhou,X. and Wilson,J.M.		
JOURNAL	Clades of Adeno-associated viruses are widely disseminated in human tissues		
PUBMED	J. Virol. 78 (12), 6381-6388 (2004)		
AUTHORS	15163731		
TITLE	2 (bases 1 to 2211)		
JOURNAL	Gao,G.; Vandenbergh,L.H., Alvira,M.R., Lu,Y., Calcedo,R., Zhou,X. and Wilson,J.M.		
FEATURES	Direct Submission		
source	Submitted (20-JAN-2004) Gene Therapy Program, Division of Medical Genetics, Department of Medicine, University of Pennsylvania, 3601 Spruce Street, Wistar Institute, Philadelphia, PA 19104, USA		
	Location/Qualifiers		
	1..2211		
	/organism="Adeno-associated virus"		
	/mol_type="genomic DNA"		
	/isolate="hu.48"		
	/isolation_source="human tissue"		
	/db_xref="taxon:272636"		
	/country="USA"		
	/notes="Clade A"		
gene	1..2211		

CDS	/gene="cap"		1. .2211		/gene="cap"		/codon_start=1		/product="capsid protein vp1"		/protein_id="AAS99296.1"		/db_xref="GI:46487869"		/translation="MAADGYLPDWLEDNLSGIREWDLKPGAPKPKANQOKQDDGRGLVLPKYTLGPFNGLDKGEVPVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEPQELRQBDTFFGNLGRAVFOAKRVLPELGLVEEGATAPGKKRPVQSPQBPDPSSGIGKTQQPAKRLNFGQTGDSVSPDPQPLGPEPPATPAAYPTTMAAGSGGAPNADNHBGADGVNAGNWHCDSTWLGDRVITSTRTWALPTYNNHLYKQISSTSTGASNDNHYFGITGNGYFDENRHFCHFSRDMWRLLNNMGFRPKRLNFKLFNIQVVEVTINOGVTITANNLTSTVQVSDSEYQLPYVLGSAHQCLPPFADPMIQQYGLYTLNNGSOAVGRSSPYCLBYFQSMRLRTGNNTFSYTFEVPFHSSYSAHSQSLDRMLNPLIDQYLYLNRQTNQSGAQNKDLLFSRSPAGMSVQPKMLPGCYRQQRVSKTKTDDNNNSNFTWTGASKYLNLRGSIINPTAVASHKDDKDFPMSGVMIQKBSAGASSATLNDNMVLTDBEEIKATNPATERFGTVAVNFQSSSTDPATGDVHAMGALPMVMQDRDVLQGPITWAKIPIHTDGHFPHSPLMGGFLGNPPQIILIKNTVPANPPAFBSATKPFASFTQYSTGQGVSEIEWELQKENSXRNWPEVQYTSNYAKSANVDTVDNNGLYTEPRPIGTRYLTRPL"	
	Query Match	99.4%;	Score	2196.2;	DB	13;	Length	2211;								
	Best Local Similarity	99.6%;	Pred.	No. 0;												
	Matches 2201;	Conservative	0;	Mismatches	8;	Indels	0;	Gaps	0;							
QY	1	ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC	60													
DB	1	ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC	60													
QY	61	GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCCAACCGCAAAAGCAGGAC	120													
DB	61	GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCCAACCGCAAAAGCAGGAC	120													
QY	121	GACGCCCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC	180													
DB	121	GACGCCCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC	180													
QY	181	AAGGGGAGCCGCTCAACCGCGCGGACGCGAGCGGCTCGAGCAGCAAGGCTTACGAC	240													
DB	181	AAGGGGAGCCGCTCAACCGCGCGGACGCGAGCGGCTCGAGCAGCAAGGCTTACGAC	240													
QY	241	CAGCAGCTCAAGCGGGTGACAATCCGTACTCGGTATAACCGCGAGCGCGAGTTT	300													
DB	241	CAGCAGCTCAAGCGGGTGACAATCCGTACTCGGTATAACCGCGAGCGCGAGTTT	300													
QY	301	CAGGAGCGTCTCAAGAGATACGTCCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG	360													
DB	301	CAGGAGCGTCTCAAGAGATACGTCCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG	360													
QY	361	GCCAAAGACGGGTTCTCGAACCTCTCGGTCTGGTTAGGAAGCGCTTAAGACGGCTCT	420													
DB	361	GCCAAAGACGGGTTCTCGAACCTCTCGGTCTGGTTAGGAAGCGCTTAAGACGGCTCT	420													
QY	421	GGAAAGAAACGTCGCGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCGGGCATTCGC	480													
DB	421	GGAAAGAAACGTCGCGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGGCATTCGC	480													
QY	481	AAGACAGGCCAGGACCGCCCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540													
DB	481	AAGACAGGCCAGGACCGCCCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540													
QY	541	TCAGTCCCGGATCGCAACCTCTCGAGAAACCTTCAGCAACCCCGCTGCTGTGGGACCT	600													
DB	541	TCAGTCCCGGATCGCAACCTCTCGAGAAACCTTCAGCAACCCCGCTGCTGTGGGACCT	600													
QY	601	ACTTCAATGGCTTCAGGCGGTGGCGACCAATGGCAGACAATAAGAGCGCGCAGCGGA	660													
DB	601	ACTTCAATGGCTTCAGGCGGTGGCGACCAATGGCAGACAATAAGAGCGCGCAGCGGA	660													
QY	661	GTGGGTAAATGCTCAGGAAATTTGGCATTCGCGATTCACATGGCTGGGCGACAGATCATC	720													
DB	661	GTGGGTAAATGCTCAGGAAATTTGGCATTCGCGATTCACATGGCTGGGCGACAGATCATC	720													


```
QY 721 ACCACGACGACCGGACCTGGGCTTGGCCACCTACAATAACCACTCTTACAAGCAATC 780
Db 721 ACCACGACGACCGGACCTGGGCTTGGCCACCTACAATAACCACTCTTACAAGCAATC 780
QY 781 TCAGTGTCTTCAACGGGGGCGAGCAACGACCACTACTTGGCTACAGCAGCCCTGG 840
Db 781 TCAGTGTCTTCAACGGGGGCGAGCAACGACCACTACTTGGCTACAGCAGCCCTGG 840
QY 841 GGTATTTTGAATTTCAACAGATTTCCACTGCGCTTTTCAACGAGTGTGGCAGGACTC 900
Db 841 GGTATTTTGAATTTCAACAGATTTCCACTGCGCTTTTCAACGAGTGTGGCAGGACTC 900
QY 901 ATCAACAACAATTTGGGGATTTCCGGCCCCAAGAGACTCAAACTTCAAACCTTTCAACATCCAA 960
Db 901 ATCAACAACAATTTGGGGATTTCCGGCCCCAAGAGACTCAAACTTCAAACCTTTCAACATCCAA 960
QY 961 GTCAAGGAGGTACGACGAATGATGGGCTTCAACCAATCGCTAATTAACCTTACCAGCAG 1020
Db 961 GTCAAGGAGGTACGACGAATGATGGGCTTCAACCAATCGCTAATTAACCTTACCAGCAG 1020
QY 1021 GTTCAAGTCTTCCGAGCTCGGAGTACCACTTCCGTAAGTCCGCTCGGCTCGGCAACAG 1080
Db 1021 GTTCAAGTCTTCCGAGCTCGGAGTACCACTTCCGTAAGTCCGCTCGGCTCGGCAACAG 1080
QY 1081 GGCTGCTTCCCTCCCTTCCGGCGGAGCTTTCATGATTCGGCAATACGCTACCTGAGC 1140
Db 1081 GGCTGCTTCCCTCCCTTCCGGCGGAGCTTTCATGATTCGGCAATACGCTACCTGAGC 1140
QY 1141 CTCAACAATGGCAGCGGAGCGTGGAGCGTTTCACTTCTTCTGCTGGAATTTTCCCT 1200
Db 1141 CTCAACAATGGCAGCGGAGCGTGGAGCGTTTCACTTCTTCTGCTGGAATTTTCCCT 1200
QY 1201 TCTCAGATCTGAGACGGGACCACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCT 1260
Db 1201 TCTCAGATCTGAGACGGGACCACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCT 1260
QY 1261 TTCCACAGCAGCTTACGCGACGACGAGCTTGGAGCGCTGATGATTCCTCTCATCGAC 1320
Db 1261 TTCCACAGCAGCTTACGCGACGACGAGCTTGGAGCGCTGATGATTCCTCTCATCGAC 1320
QY 1321 CAATACTGTATTTACCTGAGACAGAACTCAAAATCAGTTCGGAAGTCCCAAAACAAGGAC 1380
Db 1321 CAATACTGTATTTACCTGAGACAGAACTCAAAATCAGTTCGGAAGTCCCAAAACAAGGAC 1380
QY 1381 TTGCTGTTTACGCTGGCTTCCAGCTGGCATGTCTTTCAGCCCAAACTGGCTACCT 1440
Db 1381 TTGCTGTTTACGCTGGCTTCCAGCTGGCATGTCTTTCAGCCCAAACTGGCTACCT 1440
QY 1441 GGACCTGTTTATCGGACGACGCGTTCCTAAACCAAAACAGACCAACCAACAGCAAT 1500
Db 1441 GGACCTGTTTATCGGACGACGCGTTCCTAAACCAAAACAGACCAACCAACAGCAAT 1500
QY 1501 TTTTACCTGGACTGGTGTCTTCAAAATATACTTCAATGGGCGTGAATTCATCAACCT 1560
Db 1501 TTTTACCTGGACTGGTGTCTTCAAAATATACTTCAATGGGCGTGAATTCATCAACCT 1560
QY 1561 GGCACTGTCTATGGCTTACAAAGACGACGAGCAAGTTCCTTCCATGAGCGGTGTC 1620
Db 1561 GGCACTGTCTATGGCTTACAAAGACGACGAGCAAGTTCCTTCCATGAGCGGTGTC 1620
QY 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAAACTCATCTGATTTGGACAATGTCATGAT 1680
Db 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAAGCACTGATTTGGACAATGTCATGAT 1680
QY 1681 ACAGACGAAGAGGAAATTTAAAGCCACTAACCTGTGCGCCACGAAAGATTTGGGACCGTG 1740
Db 1681 ACAGACGAAGAGGAAATTTAAAGCCACTAACCTGTGCGCCACGAAAGATTTGGGACCGTG 1740
QY 1741 GCAGTCAATTTCCAGACGACGACAGACCTTCCGACCGGAGATGTGCTATGGGA 1800
Db 1741 GCAGTCAATTTCCAGACGACGACAGACCTTCCGACCGGAGATGTGCTATGGGA 1800
```

```
QY 1801 GCATTACCTGGCATGCTGTGGCAAGATAGAGACGTGTACTCCAGGGTCCCATTTGGGCG 1860
Db 1801 GCATTACCTGGCATGCTGTGGCAAGATAGAGACGTGTACTCCAGGGTCCCATTTGGGCG 1860
QY 1861 AAAATTTCTTCAACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920
Db 1861 AAAATTTCTTCAACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920
QY 1921 AAGAAACCCGCTCTCTCAGATCCTCATCAAAAAACAGCGCTGTTTCTTCCGAAATCCTCCGGG 1980
Db 1921 AAGAAACCCGCTCTCTCAGATCCTCATCAAAAAACAGCGCTGTTTCTTCCGAAATCCTCCGGG 1980
QY 1981 GAGTTTTAGCTACAAAGTTTCTTCTTATTCATCACCATACTCCACAGGACAAGTGAGT 2040
Db 1981 GAGTTTTAGCTACAAAGTTTCTTCTTATTCATCACCATACTCCACAGGACAAGTGAGT 2040
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCAAGTGACAG 2100
Db 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCAAGTGACAG 2100
QY 2101 TACACATCCAAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 2160
Db 2101 TACACATCCAAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 2160
QY 2161 TATACCTGAGCTTCCGCCCCATTTGGACCCGTTACCTTACCCGTCCTCTGT 2209
Db 2161 TATACCTGAGCTTCCGCCCCATTTGGACCCGTTACCTTACCCGTCCTCTGT 2209
```

```
RESULT 12
CS073492 LOCUS 2214 bp DNA linear PAT 05-MAY-2005
DEFINITION Sequence 80 from Patent WO2005033321.
ACCESSION CS073492
VERSION CS073492.1 GI:63090471
KEYWORDS .
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Wilson,J.M., Gao G., Alvira,M.R. and Vandenberghe,L.H.
TITLE Adeno-associated virus (aav) clades, sequences, vectors containing same, and uses thereof
JOURNAL Patent: WO 2005033321-A 80 14-APR-2005;
The Trustees of the University of Pennsylvania (US)
FEATURES
source
1..2214
/mol_type="unidentified"
/db_xref="taxon:3264"
/note="new AAV serotype, clone hu.43"
```

Query Match 96.2%; Score 2124; DB 6; Length 2214;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

```
QY 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGCAACCTCTCTGAGGSCATTTCG 60
Db 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGCAACCTCTCTGAGGSCATTTCG 60
QY 61 GAGTGTGGGACTTGAACCTTGGAGCCCGAAAGCCCAAGCCCAAGCCCAAGCAAGCAGGAC 120
Db 61 GAGTGTGGGACTTGAACCTTGGAGCCCGAAAGCCCAAGCCCAAGCCCAAGCAGGAC 120
QY 121 GACGGCCGGGCTGTGGTCTTCTTGGCTACAGTACTCGACCCCTTAAAGGACTTCGAC 180
Db 121 GACGGCCGGGCTGTGGTCTTCTTGGCTACAGTACTCGACCCCTTAAAGGACTTCGAC 180
QY 181 AAGGGGGAGCCCGTCAACCGCGGACGAGCGGCGCTCGAGCAGCAGGCTTACGAC 240
Db 181 AAGGGGGAGCCCGTCAACCGCGGACGAGCGGCGCTCGAGCAGCAGGCTTACGAC 240
```

QY 241 CAGAGCTCAAAGCGGTGACAACTCGTACCTCGGTATAACCAAGCGGACCGCGAGTTT 300
Db |||||
QY 241 CAGAGCTCAAAGCGGTGACAACTCGTACCGCGGTATAACCAAGCGGACCGCGAGTTT 300
Db |||||
QY 301 CAGGAGCGTCTGCAAGAAATACGCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
Db |||||
QY 301 CAGGAGCGTCTGCAAGAAATACGCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
Db |||||
QY 361 GCAAGAAAGCGGTCTCGAACTCTCGGTCTGGTTGAGGAAAGCGCTTAAGACGGCTCT 420
Db |||||
QY 361 GCAAGAAAGCGGTCTCGAACTCTCGGTCTGGTTGAGGAAAGCTCTTAAGACGGCTCT 420
Db |||||
QY 421 GGAAGAAAGCGTCCGGTAGAGAGTCGCCACA - - - AGAGCCAGAGCTCTCTCGGGCATC 477
Db |||||
QY 421 GGAAGAAAGAGACCGGTAGAACCGTCACTCAGCGTTCCCCCGACTCTCTCCACGGGCATC 480
Db |||||
QY 478 GGAAGAGACAGGCGAGCGCCCTGCTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCA 537
Db |||||
QY 481 GGAAGAAAGCGCGAGCGCCCTGCTAAAGAGAGACTGAACTTTGGTCAGACTGGCGACTCA 540
Db |||||
QY 538 GAGTCAGTCCCGATCCAACTCTCGGAGAACTCTCAGAGAACCCCGCGCTGTGTGGGA 597
Db |||||
QY 541 GAGTCAGTCCCGACCTCAACCAATCGGAGAACCAACAGCAGGCGCCCTCTGGTCTGGGA 600
Db |||||
QY 598 CTTACTACAAATGCTTTCAGCGGTGGCGCAACCAATGGCAGAGCAATAACGAAGCGCGAC 657
Db |||||
QY 601 TCTGGTACAAATGGCTCAGCGGTGGCGCTCCAAATGGCAGAGCAATAACGAAGCGCGAC 660
Db |||||
QY 658 GGAGTGGGTAAATGCTTCAGGAAATGGCAATTCGGAATTCACATGGTGGCGGACAGAGTC 717
Db |||||
QY 661 GGAGTGGGTAAATGCTTCAGGAAATGGCAATTCGGAATTCACATGGTGGCGGACAGAGTC 720
Db |||||
QY 718 ATCACCAAGACAGCGCACTCTGGGCTTTGCCCACTCAATAACCACTCTTACAGCAA 777
Db |||||
QY 721 ATCACCAAGACAGCGCACTCTGGGCTTTGCCCACTCAATAACCACTCTTACAGCAA 780
Db |||||
QY 778 ATCTCAGTGTCTTAACGGGGGCCAGCAAGCAACCACTACTTCGGCTACAGCAACCCC 837
Db |||||
QY 781 ATCTCAGTGTCTTAACGGGGGCCAGCAAGCAACCACTACTTCGGCTACAGCAACCCC 840
Db |||||
QY 838 TGGGGGTATTTTGAATTTCAACAGATTCCTACCTGCGCACTTTTACCACAGTACTGGAGCGA 897
Db |||||
QY 841 TGGGGGTATTTTGAATTTCAACAGATTCCTACCTGCGCACTTTTACCACAGTACTGGAGCGA 900
Db |||||
QY 898 CTCATCAACAAATTTGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC 957
Db |||||
QY 901 CTCATCAACAAATTTGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC 960
Db |||||
QY 958 CAAGTCAAGGAGTCAAGCAGATGATGGCGTCACAACCATCGCTAATAACCTTACCAGC 1017
Db |||||
QY 961 CAAGTCAAGGAGTCAAGCAGATGATGGCGTCACAACCATCGCTAATAACCTTACCAGC 1020
Db |||||
QY 1018 ACGGTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTAAGTCTCTCGGCTCTGGCAC 1077
Db |||||
QY 1021 ACGGTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTAAGTCTCTCGGCTCTGGCAC 1080
Db |||||
QY 1078 CAGGGTGTCTCTCGGTTCCGGCGGAGCGTGTTCATGATTCGCAATACGGCTACCTG 1137
Db |||||
QY 1081 CAGGGTGTCTCTCGGTTCCGGCGGAGCGTGTTCATGATTCGCAATACGGCTACCTG 1140
Db |||||
QY 1138 ACGTCAACAAATGGCAGCCAAAGCGGTGGAGCGTTCATCTTTTACTCGCTGGAATATTC 1197
Db |||||
QY 1141 ACGTCAACAAATGGCAGCCAAAGCGGTGGAGCGTTCATCTTTTACTGCTGGAATATTC 1200
Db |||||
QY 1198 CTTTCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTG 1257
Db |||||
QY 1201 CTTTCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTG 1260
Db |||||
QY 1258 CTTTCCACAGAGCTACCGGCAAGCGCAGAGCTGGACCGGCTGATGATCTCTCATC 1317
Db |||||
QY 1261 CTTTCCACAGAGCTACCGGCAAGCGCAGAGCTGGACCGGCTGATGATCTCTCATC 1320
Db |||||
QY 1318 GACCAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAGAGTCCCCAAACAG 1377
Db |||||

Db |||||
QY 1321 GTCCAAATACCTGTATTATTACCTGAACAGAACTCAAAATCAGTCGGAGTGCCCCAAACAG 1380
Db |||||
QY 1378 GACTTGTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTA 1437
Db |||||
QY 1381 GACTTGTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTA 1440
Db |||||
QY 1438 CTTGAGCCCTGTATTGGCAGAGCGGTCTTAAACAAACAAACAAACAGCAACAAACAGC 1497
Db |||||
QY 1441 CTTGAGCCCTGTATTGGCAGAGCGGTCTTAAACAAACAAACAAACAGCAACAAACAGC 1500
Db |||||
QY 1498 AATTTTACTCTGGACTGGTGTCTTCAAAATATAAACCCTCAATGGCGGTGAATTCATCATCAAC 1557
Db |||||
QY 1501 AATTTTACTCTGGACTGGTGTCTTCAAAATATAAACCCTCAATGGCGGTGAATTCATCATCAAC 1560
Db |||||
QY 1558 CTTGCACTGTCTATGGCTCTCACAAAGACAGCAAGAGTCTTCTTCCCATGAGCGGT 1617
Db |||||
QY 1561 CTTGCACTGTCTATGGCTCTCACAAAGACAGCAAGAGTCTTCTTCCCATGAGCGGT 1620
Db |||||
QY 1618 GTCATGATTTTGGAAAGAGAGCGCGGAGGTTTCAAACTGTGATTTGGACATGTCATG 1677
Db |||||
QY 1621 GTCATGATTTTGGAAAGAGAGCGCGGAGGTTTCAAACTGTGATTTGGACATGTCATG 1680
Db |||||
QY 1678 ATTACAGACGAGAGGAAATTTAAAGCCACTTAACCTGTGGCCACCAGAAAGATTTGGGACC 1737
Db |||||
QY 1681 ATTACAGACGAGAGGAAATTTAAAGCCACTTAACCTGTGGCCACCAGAAAGATTTGGGACC 1740
Db |||||
QY 1738 GTGGCAGTCAATTTTCCAGAGCAGCAGCAGACACCTTCGCGAGAGATGTGATCTATG 1797
Db |||||
QY 1741 GTGGCAGTCAATTTTCCAGAGCAGCAGCAGCAGACCTTCGCGAGAGATGTGATCTATG 1800
Db |||||
QY 1798 GGAGCATTTACTTGGCATGTGTGGCAAGATAGAGAGCTGTACTTCAGCGGTCCCATTTGG 1857
Db |||||
QY 1801 GGAGCATTTACTTGGCATGTGTGGCAAGATAGAGAGCTGTACTTCAGCGGTCCCATTTGG 1860
Db |||||
QY 1858 GCCAAAATTCCTCACACAGATGGACACTTTCACCGCTCTCTCTTATTCGGGGCTTTGGA 1917
Db |||||
QY 1861 GCCAAAATTCCTCACACAGATGGACACTTTCACCGCTCTCTCTTATTCGGGGCTTTGGA 1920
Db |||||
QY 1918 CTCAGAAACCGGCTCTCTCAGATCTCTCAATAAACAACGCTGTTCCTGCGAATCCTCG 1977
Db |||||
QY 1921 CTCAGAAACCGGCTCTCTCAGATCTCTCAATAAACAACGCTGTTCCTGCGAATCCTCG 1980
Db |||||
QY 1978 GCGGAGTTTTCAGCTACAAAGTTTGTCTTCATTCATCCCAATACTTCCACAGGACAAGTG 2037
Db |||||
QY 1981 GCGGAGTTTTCAGCTACAAAGTTTGTCTTCATTCATCCCAATACTTCCACAGGACAAGTG 2040
Db |||||
QY 2038 AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTG 2097
Db |||||
QY 2041 AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTG 2100
Db |||||
QY 2098 CAGTACATCAATTTATGCAAAATCTGCCAAAGTGTGATTTTACTGTGGACAAATGGA 2157
Db |||||
QY 2101 CAGTACATCAATTTATGCAAAATCTGCCAGCGTTGATTTTACTGTGGACAAATGGA 2160
Db |||||
QY 2158 CTTTATACTGAGCTCGGCCCATTTGGACCGGTTCACCTTACCGTCCCTGT 2209
Db |||||
QY 2161 CTTTATACTGAGCTCGGCCCATTTGGACCGGTTCACCTTACCGTCCCTGT 2212
Db |||||
RESULT 13
AY530606
LOCUS
DEFINITION
Adeno-associated virus isolate hu.43 capsid protein VP1 (cap) gene,
complete cds.
ACCESSION
AY530606
VERSION
AY530606.1 GI:46487858
KEYWORDS
Adeno-associated virus
Adeno-associated virus
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE
1 (bases 1 to 2214)
AUTHORS
Gao,G., Vandenberghe,L.H., Alvira,M.R., Lu,Y., Calcedo,R., Zhou,X.

and Wilson, J.M.
Clades of Adeno-associated viruses are widely disseminated in human tissues

J. Virol. 78 (12), 6381-6388 (2004)

15163731

2 (bases 1 to 2214)

Gao, G., Vandenbergh, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X.

and Wilson, J.M.

Direct Submission

Submitted (20-JAN-2004) Gene Therapy Program, Division of Medical

Genetics, Department of Medicine, University of Pennsylvania, 3601

Spruce Street, Wistar Institute, Philadelphia, PA 19104, USA

Location/Qualifiers

1. .2214

/organism="Adeno-associated virus"

/mol_type="genomic DNA"

/isolate="hu.43"

/isolation_source="human tissue"

/db_xref="taxon:272636"

/country="USA"

/note="clade A"

1. .2214

/gene="cap"

1. .2214

/gene="cap"

/codon_start=1

/product="capsid protein VP1"

/protein_id="AA59291.1"

/db_xref="GI:46487859"

/translation="MAADGYLPDWLEDNISEGIREWDLKPGAPKPKANQKQDDGRG

LVLPGYKILGPNGLDKGPEVNAADAALEHDKAYDQQLKAGDNPYPYRNHADAFOE

RLQEDTFPGNLGRAVFOAKRVLKGLVLEBAATKPKRPPVSPQSPDSSTGI

GKGGQPAKRLNFGTGDSEVPDQIGEPAGPSGLSGTMAAGGAPWADNNEG

ADGVGNASGHWCDLWLDGRVITSTRWALPTVNNHLYKQISSASTGSDNDHYFG

YSTPGYDFNRFHFSRDRQRLNNWGRPKRLNFKLNIQKVEVTNDGVTI

ANNLSTVQFSDSEYQVLYLGSAGHQCLPPFPADVFMIPQIYLYLNGSQAVGRS

SYFLCYFPFSSQNLNFTFYFEEVPLSHSYAHSQSLDRMLNPLIVQYLYLNR

QNSGSAONKDLLFSRGSAGMSVQPKWLPQPCYQQRVSKTKTDNNNSFTWTGAS

KYLNAGRESI INPTAMASHKODEDKFFPMGVMIFGSGASANTLONVMTDBEE

IKATNPVATERGTAVNFQSSSTDPAEDVHAMGALPGMWQDRDVLQGPIMAKIP

HTDGHFPLPMLGFGKNNPPQILIKNTIPVANPPAEFSATKPSAFITQYSTGQSVS

ETEWELQKNSKRWNEVQYTSNYAKSASVDFTVDNNGLYTBPRPIGTRYLTRPL"

ORIGIN

Query Match 96.2%; Score 2124; DB 13; Length 2214;
Best Local Similarity 97.8%; Pred. No. 0; Mismatches 45; Indels 3; Gaps 1;
Matches 2164; Conservative 0;

QY	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTCGC	60
DB	1	ATGGCTGCTGACGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTCGC	60
QY	61	GAGTGTGGGACTTGAAACCTTGAGCCCGAAGCCCAAGCCAAAGCCAGCAAGAGGAGAC	120
DB	61	GAGTGTGGGACTTGAAACCTTGAGCCCGAAGCCCAAGCCCAAGCCAGCAAGAGGAGAC	120
QY	121	GAGCGCCGGGCTCTGGTCTCTTCTGCTACAAGTACTCGGACCTTTCAACGCACTCGAC	180
DB	121	GAGCGCCGGGCTCTGGTCTCTTCTGCTACAAGTACTCGGACCTTTCAACGCACTCGAC	180
QY	181	AAGGGGAGCCCGCTCAACCGCGGCGAGCGAGCGGCCCTTCGAGCAGCAAGAGGCTACGAC	240
DB	181	AAGGGGAGCCCGCTCAACCGCGGCGAGCGAGCGGCCCTTCGAGCAGCAAGAGGCTACGAC	240
QY	241	CAGCAGCTCAAGCGGGTGACAAATCCGTACTCTCGGTTATAAACGCGCGACCGGAGTTT	300
DB	241	CAGCAGCTCAAGCGGGTGACAAATCCGTACTCTCGGTTATAAACGCGCGACCGGAGTTT	300
QY	301	CAGGAGCGCTGCAGAGAGATAGCTCTTTTGGGGGCAACCTTCGGGGGAGCAGTCTCCAG	360
DB	301	CAGGAGCGCTGCAGAGAGATAGCTCTTTTGGGGGCAACCTTCGGGGGAGCAGTCTCCAG	360
QY	361	GCCAAGAAGCGGGTTCTCGAACTCTCGGCTCGGTTGAGGAAGCGCGCTAAGACGGCTCCT	420

DB	361	GCCAAGAAGCGGGTTCTCGAACTCTCGGCTCGGTTGAGGAAGCTGCTAAGACGGCTCCT	420
QY	421	GGAAAGAAACGTCGCGGTAGAGCAGTCGCGACA---AGAGCCAGAGACTCTCTCTCGGCGATC	477
DB	421	GGAAAGAAAGACCGGTAGAACCGTCACCTCAGCGTTCCCGGACTCTCTCCACGGGCAATC	480
QY	478	GGCAAGACAGCCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTCGGCACTCA	537
DB	481	GGCAAGAAAGGCGCAGCAGCCCGCTAAAAGAGACTGAATTTGGTTCAGACTCGGCACTCA	540
QY	538	GAGTCAGTCCCGGATCCACACCTCTCGAGAGACCTCCAGCAACCCCGCTCTGTGGGA	597
DB	541	GAGTCAGTCCCGGATCCACCAATCGAGAGAACCCAGCAGAGGCCCTCTGTGTGGGA	600
QY	598	CCTACTACAATGGCTTCAGGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGAC	657
DB	601	TCTGGTACATGGCTTCAGGGCGGTGGCGCTCCAATGGCAGACAATAACGAAGCGCGAC	660
QY	658	GGAGTGGGTAAATGCTCAGGAAATTTGGCATTCGGATTCCACATGGCTGGGCGACAGATC	717
DB	661	GGAGTGGGTAAATGCTCAGGAAATTTGGCATTCGGATTCCACATGGCTGGGCGACAGATC	720
QY	718	ATCACCAACAGCACCGGACCTGGGCTTGGCCACTTACCAATAACCACTCTACAGCAA	777
DB	721	ATCACCAACAGCACCGGACCTGGGCTTGGCCACTTACCAATAACCACTCTACAGCAA	780
QY	778	ATCTCCAGTCTTCAACGGGGCCAGCAACAGCAACCACTACTTCCGGCTACAGCACCCCC	837
DB	781	ATCTCCAGTCTTCAACGGGGCCAGCAACAGCAACCACTACTTCCGGCTACAGCACCCCC	840
QY	838	TGGGGGTATTTTGAATTTCAACAGATTCCAATCCCACTTTTCAACAGCTGACGTGGCAGCGA	897
DB	841	TGGGGGTATTTTGAATTTTCAACAGATTCCAATCCCACTTTTCAACAGCTGACGTGGCAGCGA	900
QY	898	CTCATCAACAAATTTGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC	957
DB	901	CTCATCAACAAATTTGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC	960
QY	958	CAAGTCAAGGAGTCAACGAGAAATGATGGCGTCAACCACTCGCTAATAACCTTACCAGC	1017
DB	961	CAAGTCAAGGAGTCAACGAGAAATGATGGCGTCAACCACTCGCTAATAACCTTACCAGC	1020
QY	1018	ACGGTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCGGTCTCGGCTCTCGGCTAC	1077
DB	1021	ACGGTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCGGTCTCGGCTCTCGGCTAC	1080
QY	1078	CAGGGCTGCTCTCTCGGTTCCGGCGGAGCTGTTTCATGATTCGGCAATACGGCTACCTG	1137
DB	1081	CAGGGCTGCTCTCTCGGTTCCGGCGGAGCTGTTTCATGATTCGGCAATACGGCTACCTG	1140
QY	1138	ACGCTCAACAATGGCAGCCCAAGCCGTTGGGACGTTTATCTCTTTTACCTCGCTGGAATATTC	1197
DB	1141	ACGCTCAACAATGGCAGCCCAAGCCGTTGGGACGTTTATCTCTTTTACCTCGCTGGAATATTC	1200
QY	1198	CCTTCTCAGATGCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTG	1257
DB	1201	CCTTCTCAGATGCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTG	1260
QY	1258	CCTTCTCAGAGACTACCGGCAACAGGCTGGACCGGCTGATGATCTCTCTCATC	1317
DB	1261	CCTTCTCAGAGACTACCGGCAACAGGCTGGACCGGCTGATGATCTCTCTCATC	1320
QY	1318	GACCAATACCTGATTTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAAG	1377
DB	1321	GTCCAATACCTGATTTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAAG	1380
QY	1378	GACTTCTGTTTAGCGGTGGGCTCCAGCTGGCATGTCTGTTTTCAGCCCCAAAACTGGCTA	1437
DB	1381	GACTTCTGTTTAGCGGTGGGCTCCAGCTGGCATGTCTGTTTTCAGCCCCAAAACTGGCTA	1440
QY	1438	CCTTGGACCTGTTCAGGAGAGCGCGTTTCTTAAACAAAAACAGCAACAAACAGC	1497

Db	1441	CCTGGACCTGTTATCGGCAGCAGCGCGTTTCTAAACAAAAACAGACAAACAACAGC	1500	ORIGIN	
Qy	1498	AAATTACCTGACTGGTGCTTCAAAATATAAATTAACCTCAATGGGGGTGAATCCATCATCAAC	1557	Query Match	95.5%; Score 2109.8; DB 6; Length 2211;
Db	1501	AAATTACCTGACTGGTGCTTCAAAATATAAATTAACCTCAATGGGGGTGAATCCATCATCAAC	1560	Best Local Similarity	97.2%; Pred. No. 0;
Qy	1558	CCTGGCAGCTGATGCGCTCACACAAAGACGACGAAAGTTCCTTTCCCATGAGCGGT	1617	Matches 2147; Conservative	0; Mismatches 62; Indels 0; Gaps 0;
Db	1561	CCTGGCAGCTGATGCGCTCACACAAAGACGACGAAAGTTCCTTTCCCATGAGCGGT	1620		
Qy	1618	GTCAATGTTTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAATGTCATG	1677		
Db	1621	GTCAATGTTTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAATGTCATG	1680		
Qy	1678	ATTACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGAGCC	1737		
Db	1681	ATTACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGAGCC	1740		
Qy	1738	GTGGCAGTCAATTTCCAGAGCAGCAGCAGACAGACCCCTGCGACCGGAGATGTCATGATG	1797		
Db	1741	GTGGCAGTCAATTTCCAGAGCAGCAGCAGACAGACCCCTGCGACCGGAGATGTCATGATG	1800		
Qy	1798	GGAGCAATTACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGACGGGTCCCATTTGG	1857		
Db	1801	GGAGCAATTACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGACGGGTCCCATTTGG	1860		
Qy	1858	GCACAAATTCCTCACACAGATGACACTTTTCCACCGCTCTCTCTTATGGGGCGCTTTGGA	1917		
Db	1861	GCACAAATTCCTCACACAGATGACACTTTTCCACCGCTCTCTCTTATGGGGCGCTTTGGA	1920		
Qy	1918	CTCAAGAACCCGCTCCTCAGATCCCTCATCAAAACACAGCCCTGTTCCCTCGGAATCCTCG	1977		
Db	1921	CTCAAGAACCCGCTCCTCAGATCCCTCATCAAAACACAGCCCTGTTCCCTCGGAATCCTCG	1980		
Qy	1978	GGGAGTTTTTCAGTACAAAGTTTGTCTCATTCATCACCCCAATPACTCCACAGACAAAGTG	2037		
Db	1981	GGGAGTTTTTCAGTACAAAGTTTGTCTCATTCATCACCCCAATPACTCCACAGACAAAGTG	2040		
Qy	2038	AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAAAGTG	2097		
Db	2041	AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAAAGTG	2100		
Qy	2098	CAGTACACATCCAAATTCACAAATCTGCCAAGTTGATTTTACTGTGACACAAATGGA	2157		
Db	2101	CAGTACACATCCAAATTCACAAATCTGCCAAGTTGATTTTACTGTGACACAAATGGA	2160		
Qy	2158	CTTTATCTAGAGCTCGCCCCCATTTGGCACCCGTTACTCTTACCGTCCCTGT	2209		
Db	2161	CTTTATCTAGAGCTCGCCCCCATTTGGCACCCGTTACTCTTACCGTCCCTGT	2212		
RESULT 14					
LOCUS	CS073493	2211 bp	DNA	linear	PAT 05-MAY-2005
DEFINITION	Sequence 81 from Patent WO2005033321.				
ACCESSION	CS073493				
VERSION	CS073493.1	GI:63090472			
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1				
AUTHORS	Wilson, J.M.; Gao, G.; Alvira, M.R. and Vandenberghe, L.H.				
TITLE	Adeno-associated virus (aav) clades, sequences, vectors containing				
JOURNAL	same, and uses thereof				
Patent:	WO 2005033321-A 81 14-APR-2005;				
The Trustees of the University of Pennsylvania (US)					
FEATURES	Location/Qualifiers				
source	1..2211				
	/organism="unidentified"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32644"				
	/note="new AAV serotype, clone hu.44"				

```
QY 1021 GTTCAAGTCTTTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTCTGGCCACCCAG 1080
Db 1021 GTTCAAGTCTTTCGGACTCGGAGTACAGCTTCCGTACGTCTCGGCTCTGGCCACCCAG 1080
QY 1081 GGTGCTCTCCCTCGGTTCCGGCGGAGCGTGTTCATGATTCGGCAATACGGCTACCTGACG 1140
Db 1081 GGTGCTCTCCCTCGGTTCCGGCGGAGCGTGTTCATGATTCGGCAATACGGCTACCTGACG 1140
QY 1141 CTCAACAATGGCAGCAAGCGGTGGAGCGTTTCATCTCTTTTACTGCTTGGCAATATTTCCCT 1200
Db 1141 CTCAACAATGGCAGCAAGCGGTGGAGCGTTTCATCTCTTTTACTGCTTGGCAATATTTCCCT 1200
QY 1201 TCTCAGATGCTGAGAACCGGCAACAACCTTTTACCTTTCAGCTACACCTTTTGGAGGAAGTGCCT 1260
Db 1201 TCTCAGATGCTGAGAACCGGCAACAACCTTTTACCTTTCAGCTACACCTTTTGGAGGAAGTGCCT 1260
QY 1261 TTCACAGCAGCTACCGGCAACAGCAGCGCTCGACCGGTGATGATTCCTCTCATCGAC 1320
Db 1261 TTCACAGCAGCTACCGGCAACAGCAGCGCTCGACCGGTGATGATTCCTCTCATCGAC 1320
QY 1321 CAATACCTGATTTACCTGACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC 1380
Db 1321 CAATACCTGATTTACCTGACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC 1380
QY 1381 TTGCTGTTTACGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGGCCAAAACTGGCTACCT 1440
Db 1381 TTGCTGTTTACGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGGCCAAAACTGGCTACCT 1440
QY 1441 GSAACCTGTTTATCGGACGAGCGGTTTCTTAAACAAAAACAGACAAACAACAGCAAT 1500
Db 1441 GSAACCTGTTTATCGGACGAGCGGTTTCTTAAACAAAAACAGACAAACAACAGCAAT 1500
QY 1501 TTTACCTGAGTGGTCTTCAAAATATAAATCAATGAGGCGTCAATCCATCATCAACCT 1560
Db 1501 TTTACCTGAGTGGTCTTCAAAATATAAATCAATGAGGCGTCAATCCATCATCAACCT 1560
QY 1561 GGCACCTGATGGCTCACAAAGACGACGAGCAAGCAAGTCTTTCCATGAGCGGTGTC 1620
Db 1561 GGCACCTGATGGCTCACAAAGACGACGAGCAAGTCTTTCCATGAGCGGTGTC 1620
QY 1621 ATGATTTTGGAAAGAGCGCGGAGCTTCAAACTGATGGAATGGAATGGAATGGAAT 1680
Db 1621 ATGATTTTGGAAAGAGCGCGGAGCTTCAAACTGATGGAATGGAATGGAATGGAAT 1680
QY 1681 ACAGACGAGGAAATTAAGCCACTAAACCTGCTGCGCCACCGAAGATTTGGGACCGTG 1740
Db 1681 ACAGACGAGGAAATTAAGCCACTAAACCTGCTGCGCCACCGAAGATTTGGGACCGTG 1740
QY 1741 GCAGTCAATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTCATGCGGA 1800
Db 1741 GCAGTCAATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTCATGCGGA 1800
QY 1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCGAGGTCCTCATTTGGGCC 1860
Db 1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCGAGGTCCTCATTTGGGCC 1860
QY 1861 AAAATTCCTCACAGATGAGACCTTCAACCCGTCCTCTTATGCGCGGCTTTGGACTC 1920
Db 1861 AAAATTCCTCACAGATGAGACCTTCAACCCGTCCTCTTATGCGCGGCTTTGGACTC 1920
QY 1921 AAGAACCCGCTCTCAGATCTCATCAAAACACAGCGCTGTCTCTGCGAATCTCTCGCGCG 1980
Db 1921 AAGAACCCGCTCTCAGATCTCATCAAAACACAGCGCTGTCTCTGCGAATCTCTCGCGCG 1980
QY 1981 GAGTTTTCAGTACAAAGTTTGTCTTATTCATCAACCACTCCACAGACAAAGTCAAT 2040
Db 1981 GAGTTTTCAGTACAAAGTTTGTCTTATTCATCAACCACTCCACAGACAAAGTCAAT 2040
QY 2041 GTGAAATTTGAATGGGAGCTGCAGAAAGAAACAGAGAGCGCTGGAATCCGGAAGTGCAG 2100
Db 2041 GTGAAATTTGAATGGGAGCTGCAGAAAGAAACAGAGAGCGCTGGAATCCGGAAGTGCAG 2100
```

```
QY 2101 TACATCATCAATATGCAAAATCTGCCAAGCTTGATTTTACTGTGGACAAATGACACTT 2160
Db 2101 TACATCATCAATATGCAAAATCTGCCAAGCTTGATTTTACTGTGGACAAATGACACTT 2160
QY 2161 TATACTGAGCCTCGCCCATTCGCACCGCTTACCTTACCGTCCCTCTGT 2209
Db 2161 TATACTGAGCCTCGCCCATTCGCACCGCTTACCTTACCGTCCCTCTGT 2209

RESULT 15
AY530607 2211 bp DNA linear VRL 24-JUN-2004
LOCUS Adeno-associated virus isolate hu.44 capsid protein VP1 (cap) gene,
DEFINITION complete cds.
ACCESSION AY530607
VERSION AY530607.1 GI:46487860
KEYWORDS Adeno-associated virus
SOURCE Adeno-associated virus
ORGANISM Adeno-associated virus
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 1 to 2211)
AUTHORS Gao, G., Vandenbergh, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X.
and Wilson, J.M.
TITLES Clades of Adeno-associated viruses are widely disseminated in human
tissues
JOURNAL J. Virol. 78 (12), 6381-6388 (2004)
PUBMED 15163731
REFERENCE 2 (bases 1 to 2211)
AUTHORS Gao, G., Vandenbergh, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X.
and Wilson, J.M.
TITLES Direct Submission
Submitted (20-JAN-2004) Gene Therapy Program, Division of Medical
Genetics, Department of Medicine, University of Pennsylvania, 3601
Spruce Street, Wistar Institute, Philadelphia, PA 19104, USA
FEATURES
Location/Qualifiers
1..2211
/organism="Adeno-associated virus"
/mol_type="genomic DNA"
/isolates="hu.44"
/isolation_source="human tissue"
/db_xref="taxon:272636"
/country="USA"
/notes="Clade A"
1..2211
/gene="cap"
1..2211
/gene="cap"
/codon_start=1
/product="capsid protein VP1"
/protein_id="AAS9292.1"
/db_xref="GI:46487861"
/tranlation="MAADGVLPMLEDTLSEGIROWKLRPGPPPKPABRHKDDSRG
LVLPGYKLGPFNGLDKGPEVNEADAALEHDKAYDRQDSDGNPVLKYNHDAEFOE
RLKEDTSGFNLRGAVFQAKRLEFLGVEGATAPGKRVEQSPQGPDSGGIG
KTGQPAKRLNFGDTSVPDQPLGEPATPAAGVTTWASGGGAPMADNNEGA
DVGQASGNHCDSTWLGRVITSTRWALPTNNHLYKQISSASTGSNDHYFGY
SPWGYDFNRFHCPSPRDWORLNNNGFRPKLNFKNLQVKEVTNDGVTITTA
NNLTSTVOVPSDSEYOLPVVLSAHOGLPPRPADVFMIPQVYLTLNNGSQAVGRSS
FYCLEYFPQMLATGNFTFSTFEVPHSSVYAHQSOLDRLANPLIDOLYLYPNRTQ
NOSGSAQNKLIFSRGSPAGMSVQPKNWLPGPCTRQQRVSKTKTDNNNSFTWTGASK
YNLNGRESIINPTAMASHKDDKEDFFPMSGVMI FKESAGASATNDLNNMTDDEBI
KATNPVATFRFGTVAVNFQSSSTDPATGDVHAMGALPGVMQGRDYLQGPVWAKIPH
TDGHPHPSPLMGFGFLKNPQILIKNTPVPANPPAEFSATKEPASFITQYSTQGVSE
IEWELQKENSKRNPVEQVTSYNAKSANVDFTVDNNGLYTEPRPIGTRYLTRL"

Query Match 95.5%; Score 2109.8; DB 13; Length 2211;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1 ATGCGTCCCATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCATTCGC 60
Db 1 ATGCGTCCCATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCATTCGA 60
```

QY	61	GAGTGGTGGGACTTGAACCTCGAGCCCGAAGCCCAAGCCAAAGCCAAAGCAGCAAGCAGGAC	120
DB	61	CAGTGGTGAAGCTCAGACTCGGCCCAACCAACCAAGCCCGCAGAGCGGATAAGGAC	120
QY	121	GACGGCCGGGCTCTGGTCTCTCGCTACAAAGTACTCGGACCCCTTCAACGGACTCGAC	180
DB	121	GACGAGGGGCTCTGGTCTCTCGCTACAAAGTACTCGGACCCCTTCAACGGACTCGAC	180
QY	181	AAGGGGAGCCCGTCAACCGCGCGAGCGCGGCCCTCGAGCAGCAAAAGCCTACGAC	240
DB	181	AAGGGAGAGCCCGTCAACCGCGCGAGCGCGGCCCTCGAGCAGCAAAAGCCTACGAC	240
QY	241	CAGGAGCTCAAGCGGGTACAACTCGGTACTCGGTATAACCAAGCGCGAGCGGATTT	300
DB	241	CGGAGCTCGACAGCGGAGACAACCGGTACTCAAGTACAACCAAGCGCGAGCGGATTT	300
QY	301	CAGGAGCTCTGCAAGAGATACGTCTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
DB	301	CAGGAGCGCTTAAGAGATACGTCTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
QY	361	GCCAGAGCGGGTCTCGAACTCTCGTCTGGTTGAGGAAGCGCTTAAGACGGTCTCT	420
DB	361	GCCAGAGCGGGTCTCGAACTCTCGTCTGGTTGAGGAAGCGCTTAAGACGGTCTCT	420
QY	421	GGAGAGAAAGCTCCGGTAGAGCAGTCCGACAGAGCAGACTCTCTCGGGCATCGGC	480
DB	421	GGAGAGAAAGCTCCGGTAGAGCAGTCCGACAGAGCAGACTCTCTCGGGCATCGGC	480
QY	481	AAGACAGGCGCAGCAGCCCGCTTAAAGAGACTCAATTTTGGTTCAGACTCGGACTCAGAG	540
DB	481	AAGACAGGCGCAGCAGCCCGCTTAAAGAGACTCAATTTTGGTTCAGACTCGGACTCAGAG	540
QY	541	TCAGTCCCAGATCCCAACCTCTCGGAGAACTTCCAGCAACCCCGCTCTGTGGGACCT	600
DB	541	TCAGTCCCAGATCCCAACCTCTCGGAGAACTTCCAGCAACCCCGCTCTGTGGGACCT	600
QY	601	ACTAAGATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAAGCGCGCAGCA	660
DB	601	ACTAAGATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAAGCGCGCAGCA	660
QY	661	GTGGGTAAATGCTTCAGAAATGGCATTCGATTCACATGGCTGGGCGACAGAGTCATC	720
DB	661	GTGGGTAAATGCTTCAGAAATGGCATTCGATTCACATGGCTGGGCGACAGAGTCATC	720
QY	721	ACCACAGCACCGGACCTTGGGCTTGGCCACCTTACAATAACCACTCTACAAGCAAAATC	780
DB	721	ACCACAGCACCGGACCTTGGGCTTGGCCACCTTACAATAACCACTCTACAAGCAAAATC	780
QY	781	TCCAGTGCTTCAACGGGGCCAGCAACGACCAACCACTACTTGGCTACAGCAACCCCTGG	840
DB	781	TCCAGTGCTTCAACGGGGCCAGCAACGACCAACCACTACTTGGCTACAGCAACCCCTGG	840
QY	841	GGGTATTTGATTTCAACAGATTCACCTGCACTTTTCAACAGCTGCTGGCAGGACTC	900
DB	841	GGGTATTTGATTTCAACAGATTCACCTGCACTTTTCAACAGCTGCTGGCAGGACTC	900
QY	901	ATCAACAACTTTGGGATTTCCGGCCCAAGAGACTCAAATTCACCTTCAACATCCAA	960
DB	901	ATCAACAACTTTGGGATTTCCGGCCCAAGAGACTCAAATTCACCTTCAACATCCAA	960
QY	961	GTCAAGAGGTACAGCAAGATGATGGGTACAGCAACCACTCGCTTAACCTTACCAGCAG	1020
DB	961	GTCAAGAGGTACAGCAAGATGATGGGTACAGCAACCACTCGCTTAACCTTACCAGCAG	1020
QY	1021	GTTCAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGTCTCGGCTCTGGCAGCAG	1080
DB	1021	GTTCAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGTCTCGGCTCTGGCAGCAG	1080
QY	1081	GGTGTCTCTCTCGGCTTCCGGCGGACGTGTTCATGATTCGCAATACGGCTACCTGACG	1140
DB	1081	GGTGTCTCTCTCGGCTTCCGGCGGACGTGTTCATGATTCGCAATACGGCTACCTGACG	1140

QY	1141	CTCAACATGACGACCGCGTGGAGTTCATCTCTTTTACTGCTCGGAATATTTCCCT	1200
DB	1141	CTCAACATGACGACCGCGTGGAGTTCATCTCTTTTACTGCTCGGAATATTTCCCT	1200
QY	1201	TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCT	1260
DB	1201	TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGCCT	1260
QY	1261	TTTCAACAGCAGCTACGCGCACAGCAGCCTTGAACCGGCTGATGAATCTCTCATCGAC	1320
DB	1261	TTTCAACAGCAGCTACGCGCACAGCAGCCTTGAACCGGCTGATGAATCTCTCATCGAC	1320
QY	1321	CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAGGAC	1380
DB	1321	CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAGGAC	1380
QY	1381	TTTGTCTTTAGCTGGGTCTTCCAGCTGGCATGTCTTTCAGCCCAAACTGGCTACCT	1440
DB	1381	TTTGTCTTTAGCTGGGTCTTCCAGCTGGCATGTCTTTCAGCCCAAACTGGCTACCT	1440
QY	1441	GGACCTCTGTATTCGGCAGCAGCGGCTTTCTAAAACAAAACAGAAACAGAAACAGCAAT	1500
DB	1441	GGACCTCTGTATTCGGCAGCAGCGGCTTTCTAAAACAAAACAGAAACAGAAACAGCAAT	1500
QY	1501	TTTACCTCGACTGGTCTTCAAAATATAAATCAATCTCAATGGGCGTGAATCATCAACCT	1560
DB	1501	TTTACCTCGACTGGTCTTCAAAATATAAATCAATCTCAATGGGCGTGAATCATCAACCT	1560
QY	1561	GGCAGTCTATGGCTCTCACAAAACAGCAGCGGCTTCAAAAGTCTTTTCCCATGAGCGGTGC	1620
DB	1561	GGCAGTCTATGGCTCTCACAAAACAGCAGCGGCTTCAAAAGTCTTTTCCCATGAGCGGTGC	1620
QY	1621	ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACACTGCAATTGGGAACAATGTGATTT	1680
DB	1621	ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACACTGCAATTGGGAACAATGTGATTT	1680
QY	1681	ACAGACGAGAGGAAATTAAGCCACTAACTCTGGCCACCGAAGATTTGGGACCGTG	1740
DB	1681	ACAGACGAGAGGAAATTAAGCCACTAACTCTGGCCACCGAAGATTTGGGACCGTG	1740
QY	1741	GCAGTCAATTTTCAGAGCAGCAGCAGACAGCCCTGCGACCGGAGATGTGATGATGGGA	1800
DB	1741	GCAGTCAATTTTCAGAGCAGCAGCAGACAGCCCTGCGACCGGAGATGTGATGATGGGA	1800
QY	1801	GCAATACCTGGCATGGTGTGGCAAGATAGAGAGCTGTAACCTGCAAGGCTCCCATTTGGGCC	1860
DB	1801	GCAATACCTGGCATGGTGTGGCAAGATAGAGAGCTGTAACCTGCAAGGCTCCCATTTGGGCC	1860
QY	1861	AAAAATTCCTCACACAGATGGACACTTTTCAACCGCTCTCTTTATGGGCGGCTTTGGACTC	1920
DB	1861	AAAAATTCCTCACACAGATGGACACTTTTCAACCGCTCTCTTTATGGGCGGCTTTGGACTC	1920
QY	1921	AAGAACCCGCTCTCAGATCCTCATAAAAACAGCCCTGTTTCTTCCGAAATCTCCGGCG	1980
DB	1921	AAGAACCCGCTCTCAGATCCTCATAAAAACAGCCCTGTTTCTTCCGAAATCTCCGGCG	1980
QY	1981	GAGTTTTTTCAGCTACAAAGTTTGTCTTATTCATCCCAATCTCCACAGGACAAGTGA	2040
DB	1981	GAGTTTTTTCAGCTACAAAGTTTGTCTTATTCATCCCAATCTCCACAGGACAAGTGA	2040
QY	2041	GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTCGGAATCCCGAAGTGCAG	2100
DB	2041	GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTCGGAATCCCGAAGTGCAG	2100
QY	2101	TACACATCCAAATTTATGCAAAATCTGCCAAACGTTTGTATTTTACTGTGGACAACAAATG	2160
DB	2101	TACACATCCAAATTTATGCAAAATCTGCCAAACGTTTGTATTTTACTGTGGACAACAAATG	2160
QY	2161	TATATGAGGCTCGGCCCAATTTGGGACCCGTTTACCTTACCTTACCGTCCCTGT	2209
DB	2161	TATATGAGGCTCGGCCCAATTTGGGACCCGTTTACCTTACCTTACCGTCCCTGT	2209

Search completed: November 29, 2005, 00:05:32
Job time : 8078.36 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:15:02 ; Search time 20.3389 Seconds
(without alignments)
12497.813 Million cell updates/sec

Title: US-10-696-900-1_COPY_1_143

Perfect score: 143

Sequence: 1 ttgcccactccctctctgcg.....aactccatcactagggttaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
- 7: /cgn2_6/prodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/prodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/prodata/1/ina/backfilese1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	4718	3	US-09-807-802A-1
2	137.2	95.9	4683	3	US-09-807-802A-19
3	119	83.2	4718	3	US-09-807-802A-1
4	107.4	75.1	145	2	US-07-789-917A-1
5	107.4	75.1	145	3	US-08-702-573-4
6	107.4	75.1	145	3	US-07-982-193-1
7	107.4	75.1	145	3	US-09-782-378A-6
8	107.4	75.1	165	2	US-07-989-841A-1
9	107.4	75.1	165	2	US-08-440-738A-1
10	107.4	75.1	165	3	US-08-471-914-1
11	107.4	75.1	165	3	US-09-276-625-7
12	107.4	75.1	165	3	US-10-054-665A-7
13	107.4	75.1	165	3	US-09-782-378A-8
14	107.4	75.1	212	3	US-09-276-625-6
15	107.4	75.1	212	3	US-10-054-665A-6
16	107.4	75.1	4675	3	US-09-782-378A-1
17	107.4	75.1	4675	3	US-09-782-378A-2
18	107.4	75.1	4679	3	US-10-038-972A-12
19	107.4	75.1	4680	2	US-08-254-358-1
20	107.4	75.1	4680	2	US-08-475-391-1
21	107.4	75.1	4680	2	US-08-709-609-1
22	107.4	75.1	4680	6	PCT-US95-07178-1
23	107.4	75.1	4681	3	US-09-807-802A-18
24	107.4	75.1	4683	3	US-09-807-802A-19

25	107.4	75.1	5932	3	US-09-299-141-4	Sequence 4, Appli
26	107.4	75.1	5932	3	US-09-299-141-4	Sequence 4, Appli
27	107.4	75.1	6142	3	US-09-299-141-8	Sequence 8, Appli
28	107.4	75.1	6142	3	US-09-299-141-8	Sequence 8, Appli
29	107.4	75.1	6253	3	US-08-893-327-15	Sequence 15, Appli
30	107.4	75.1	6253	3	US-08-893-327-15	Sequence 15, Appli
31	107.4	75.1	6280	3	US-08-893-327-17	Sequence 17, Appli
32	107.4	75.1	6280	3	US-08-893-327-17	Sequence 17, Appli
33	107.4	75.1	6280	3	US-08-893-327-19	Sequence 19, Appli
34	107.4	75.1	6280	3	US-08-893-327-19	Sequence 19, Appli
35	107.4	75.1	6565	3	US-09-299-141-1	Sequence 1, Appli
36	107.4	75.1	6565	3	US-09-299-141-1	Sequence 1, Appli
37	107.4	75.1	6714	3	US-09-299-141-6	Sequence 6, Appli
38	107.4	75.1	6714	3	US-09-299-141-6	Sequence 6, Appli
39	107.4	75.1	6924	3	US-09-299-141-9	Sequence 9, Appli
40	107.4	75.1	6924	3	US-09-299-141-9	Sequence 9, Appli
41	107.4	75.1	6924	3	US-09-299-141-10	Sequence 10, Appli
42	107.4	75.1	6924	3	US-09-299-141-10	Sequence 10, Appli
43	107.4	75.1	6924	3	US-09-299-141-11	Sequence 11, Appli
44	107.4	75.1	6924	3	US-09-299-141-11	Sequence 11, Appli
45	107.4	75.1	6981	3	US-09-299-141-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-807-802A-1
; Sequence 1, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVEN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:
US-09-807-802A-1

Query Match 100.0%; Score 143; DB 3; Length 4718;

Best Local Similarity 100.0%; Pred. No. 2.7e-30; Indels 0; Gaps 0;
Matches 143; Conservative 0; Mismatches 0;

QY	1	TTGCCACATCCCTCTCTCGCGCGCTCGCTCGGTGGGGCCCTGCGGACCAAAAGGTCCCGC	60
Db	1	TTGCCACATCCCTCTCTCGCGCGCTCGCTCGGTGGGGCCCTGCGGACCAAAAGGTCCCGC	60
QY	61	AGACGGCAGAGCTCTCTCTGCGGGCCCAACCGAGCGAGCGCGCAGAGGAGGTG	120
Db	61	AGACGGCAGAGCTCTCTCTGCGGGCCCAACCGAGCGAGCGCGCAGAGGAGGTG	120
QY	121	GGCACTCCATCACTAGGGGTAA	143
Db	121	GGCACTCCATCACTAGGGGTAA	143

RESULT 2
US-09-807-802A-19/c
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: Vectors and Host Cells Containing Same
; CURRENT APPLICATION NUMBER: GNPVN.031USA
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19

Query Match 95.9%; Score 137.2; DB 3; Length 4683;
Best Local Similarity 97.9%; Pred. No. 1.1e-28;
Matches 139; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 60
Db 4683 TTGGCCACTCCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 4624
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 4623 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 4564
QY 121 GGCACTCCATCACTAGGGGTA 142
Db 4563 GGCACTCCATCACTAGGGGTA 4542

RESULT 3
US-09-807-802A-1/c
; Sequence 1, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE REFERENCE: Vectors and Host Cells Containing Same
; CURRENT APPLICATION NUMBER: GNPVN.031USA
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:

US-09-807-802A-1

Query Match 83.2%; Score 119; DB 3; Length 4718;
Best Local Similarity 89.5%; Pred. No. 1.1e-23;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 60
Db 4718 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 4659
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 4658 CGTCTGGGAGCTTTGGTCCGAGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599
QY 121 GGCACTCCATCACTAGGGGTA 143
Db 4598 GGCACTCCATCACTAGGGGTA 4576

RESULT 4

US-07-789-917A-1
; Sequence 1, Application US/07789917A
; Patent No. 5252479
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release q.0, Version q.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,917A
; FILING DATE: 19911118
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William B.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-789-917A-1

Query Match 75.1%; Score 107.4; DB 2; Length 145;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGGGCGGACCAAAGGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCGGCGGCTCAGTGAGCGAGCGCGCAGAGAGGAGTG 120
QY 121 GGCACTCCATCACTAGGGGT 141

Db 121 GCCAACTCCATCACTAGGGT 141

RESULT 5

US-08-702-573-4

Sequence 4, Application US/08702573

Patent No. 6033885

GENERAL INFORMATION:

APPLICANT: LATTA, Martine

APPLICANT: DENEPIE, Patrice

APPLICANT: VIGNE, Emmanuelle

APPLICANT: PERRICAUDET, Michel

TITLE OF INVENTION: INTEGRATION RECOMBINANT ADENOVIRUSES,

TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.

STREET: 500 Arcola Rd. 3C43

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/702,573

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 94/02445

FILING DATE: 03-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/FR95/00233

FILING DATE: 28-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: 38,619

REFERENCE/DOCKET NUMBER: ST94011-US

TELEPHONE: (610)454-3839

TELEFAX: (610)454-3808

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 145 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..145

OTHER INFORMATION: /note= "Minimal ITR Sequence"

US-08-702-573-4

Query Match 75.1%; Score 107.4; DB 3; Length 145;

Best Local Similarity 85.1%; Pred. No. 1.2e-20;

Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAAGGTCCGC 60

Db 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAAGGTCCGC 60

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGGACGAGAGGGAGTG 120

Db 61 CGACGCCCGGGCTTTGCCCCGGGGCTCTAGTGAGCGAGCGCGGACGAGAGGGAGTG 120

QY 121 GGCAACTCCATCACTAGGGT 141

Db 121 GGCAACTCCATCACTAGGGT 141

RESULT 6

US-07-982-193-1

Sequence 1, Application US/07982193

Patent No. 6261834

GENERAL INFORMATION:

APPLICANT: Srivastava, Arun

TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/982,193

FILING DATE: 19921125

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McNulty, William E.

REGISTRATION NUMBER: 22,606

REFERENCE/DOCKET NUMBER: 8361

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 145 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-07-982-193-1

Query Match 75.1%; Score 107.4; DB 3; Length 145;

Best Local Similarity 85.1%; Pred. No. 1.2e-20;

Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAAGGTCCGC 60

Db 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAAGGTCCGC 60

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGGACGAGAGGGAGTG 120

Db 61 CGACGCCCGGGCTTTGCCCCGGGGCTCTAGTGAGCGAGCGCGGACGAGAGGGAGTG 120

QY 121 GGCAACTCCATCACTAGGGT 141

Db 121 GGCAACTCCATCACTAGGGT 141

RESULT 7

US-09-782-378A-6

Sequence 6, Application US/09782378A

Patent No. 6916635

GENERAL INFORMATION:

APPLICANT: Hearing, Patrick

APPLICANT: Bahou, Wadie

APPLICANT: Sandalon, Ziv

APPLICANT: Gnatenko, Dmitri

TITLE OF INVENTION: Adenoviral Vectors

FILE REFERENCE: STONYB-04970

CURRENT APPLICATION NUMBER: US/09/782,378A

CURRENT FILING DATE: 2001-02-12

```
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-6

Query Match      75.1%; Score 107.4; DB 3; Length 145;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAGGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAGGTCCGC 80
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCGCAGAGAGGAGTG 120
Db 81 CGACGCCGGGCTTTGCCCGGCGGCTCAGTGAGCGAGCGCGCGCAGAGAGGAGTG 140
QY 121 GGCAACTCCATCACTAGGGGT 141
Db 141 GGCAACTCCATCACTAGGGGT 161

RESULT 9
US-08-440-738A-1
; Sequence 1, Application US/08440738A
; Patent No. 5869305
; GENERAL INFORMATION:
; APPLICANT: Samuleki, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,738A
; FILING DATE: May 15, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-440-738A-1

Query Match      75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAGGTCCGC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAGGTCCGC 80
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCGCAGAGAGGAGTG 120
Db 81 CGACGCCGGGCTTTGCCCGGCGGCTCAGTGAGCGAGCGCGCGCAGAGAGGAGTG 140
QY 121 GGCAACTCCATCACTAGGGGT 141
Db 141 GGCAACTCCATCACTAGGGGT 161

RESULT 10
US-07-989-841A-1
; Sequence 1, Application US/07989841A
; Patent No. 5478745
; GENERAL INFORMATION:
; APPLICANT: Samuleki, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,841A
; FILING DATE: On even date herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-07-989-841A-1

Query Match      75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```

```

Db      81  CGAGCCCGGGCTTTGCCCGGGCGCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTG 140
QY      121  GGCAACTCCATCACTAGGGGT 141
      |||||
Db      141  GCCAACTCCATCACTAGGGGT 161

RESULT 12
US-10-054-665A-7
; Sequence 7, Application US/10054665A
; Patent No. 6897045
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongshen
; TITLE OF INVENTION: University of Iowa Research Foundation
; FILE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054.665A
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 09/276,625
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6897045 5,478,745
US-10-054-665A-7

Query Match      75.1%; Score 107.4; DB 3; Length 165;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0

QY      1  TTGCCCATCTCCCTCTCTCGGGCGCTCGCTCGCTCGGTGGGGCTCGGACACCAAGGTCGC 60
Db      21  TTGGCCACTCCCTCTCTCGGGCGCTCGCTCGCTCGTCTACTGAGCGCGGGCGACCAAGGTCGCC 80

QY      61  AGACGCGAGACTCTGCTCTCGCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 120
Db      81  CGAGCCCGGGCTTTGCCCGGGCGCTCAGTGAGCGAGCGCGCGCAGAGGGAGTG 140

QY      121  GGCAACTCCATCACTAGGGGT 141
      |||||
Db      141  GCCAACTCCATCACTAGGGGT 161

RESULT 13
US-09-782-378A-8
; Sequence 8, Application US/09782378A
; Patent No. 6916635
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782.378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-8

```

Query Match 75.1%; Score 107.4; DB 3; Length 165;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCCAAAGGTCCGC 60
DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGGTCCGC 80
QY 61 AGACGGCAGAGTCTGCTTGTGCGGCCCCACCGAGGAGCGGCGGAGAGGAGGTG 120
DB 81 CGACGCCCGGGCTTTGGTGGCGCGGCTCAGTGAGCGAGCGGCGGAGAGGAGGTG 140
QY 121 GGCAACTCCATCACTAGGGGT 141
DB 141 GGCAACTCCATCACTAGGGGT 161

RESULT 14
US-09-276-625-6
; Sequence 6, Application US/09276625
; Patent No. 6436392
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US1
; CURRENT APPLICATION NUMBER: US/09/276,625
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; TYPE: DNA
; ORGANISM: AAV circular intermediate, clone p1202
US-09-276-625-6

Query Match 75.1%; Score 107.4; DB 3; Length 272;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCCAAAGGTCCGC 60
DB 69 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCCAAAGGTCCGC 128
QY 61 AGACGGCAGAGTCTGCTCTGCGGCCCCACCGAGGAGCGGCGGAGAGGAGGTG 120
DB 129 CGACGCCCGGGCTTTGGTGGCGCGGCTCAGTGAGCGAGCGGCGGAGAGGAGGTG 188
QY 121 GGCAACTCCATCACTAGGGGT 141
DB 189 GGCAACTCCATCACTAGGGGT 209

RESULT 15
US-10-054-665A-6
; Sequence 6, Application US/10054665A
; Patent No. 6897045
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongshen
; TITLE OF INVENTION: University of Iowa Research Foundation
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054,665A
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 09/276,625
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6
; LENGTH: 272
; TYPE: DNA
; ORGANISM: AAV circular intermediate, clone p1202
US-10-054-665A-6
Query Match 75.1%; Score 107.4; DB 3; Length 272;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCCAAAGGTCCGC 60
DB 69 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCCAAAGGTCCGC 128
QY 61 AGACGGCAGAGTCTGCTCTGCGGCCCCACCGAGGAGCGGCGGAGAGGAGGTG 120
DB 129 CGACGCCCGGGCTTTGGTGGCGCGGCTCAGTGAGCGAGCGGCGGAGAGGAGGTG 188
QY 121 GGCAACTCCATCACTAGGGGT 141
DB 189 GGCAACTCCATCACTAGGGGT 209

Search completed: November 29, 2005, 14:27:39
Job time : 21.3389 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:15:02 ; Search time 20.3389 Seconds
(without alignments)
12497.813 Million cell updates/sec

Title: US-10-696-900-1_COPY_4576_4718

Perfect score: 143

Sequence: 1 ttacccttagtgatggagtt.....cgccagagaggagtgaggcaaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1 COMB.seq.*

2: /cgn2_6/prodata/1/ina/5 COMB.seq.*

3: /cgn2_6/prodata/1/ina/6A COMB.seq.*

4: /cgn2_6/prodata/1/ina/6B COMB.seq.*

5: /cgn2_6/prodata/1/ina/H COMB.seq.*

6: /cgn2_6/prodata/1/ina/PCBUS COMB.seq.*

7: /cgn2_6/prodata/1/ina/PP COMB.seq.*

8: /cgn2_6/prodata/1/ina/RE COMB.seq.*

9: /cgn2_6/prodata/1/ina/backfileseq1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	4718	3	US-09-807-802A-1
2	120.2	84.1	4683	3	US-09-807-802A-19
3	119	83.2	4718	3	US-09-807-802A-1
4	113.2	79.2	4683	3	US-09-807-802A-19
5	109	76.2	4681	3	US-09-807-802A-18
6	107.4	75.1	145	3	US-08-525-866-1
7	107.4	75.1	165	2	US-07-989-841A-1
8	107.4	75.1	165	2	US-08-440-738A-1
9	107.4	75.1	165	3	US-08-471-914-1
10	107.4	75.1	165	3	US-09-276-625-7
11	107.4	75.1	165	3	US-10-054-665A-7
12	107.4	75.1	192	3	US-08-702-573-3
13	107.4	75.1	272	3	US-09-276-625-6
14	107.4	75.1	272	3	US-10-054-665A-6
15	107.4	75.1	4675	3	US-09-782-378A-1
16	107.4	75.1	4675	3	US-09-782-378A-2
17	107.4	75.1	4679	3	US-10-038-972A-12
18	107.4	75.1	4680	2	US-08-254-358-1
19	107.4	75.1	4680	2	US-08-475-391-1
20	107.4	75.1	4680	2	US-08-709-609-1
21	107.4	75.1	4680	6	PCT-US95-07178-1
22	107.4	75.1	8698	3	US-09-770-315-2
23	105.8	74.0	272	3	US-09-276-625-4
24	105.8	74.0	272	3	US-10-054-665A-4

c 25	105.4	73.7	174	3	US-09-394-110A-1	Sequence 1, Appli
c 26	102.6	71.7	272	3	US-09-276-625-6	Sequence 6, Appli
c 27	102.6	71.7	272	3	US-10-054-665A-6	Sequence 6, Appli
c 28	101	70.6	183	3	US-09-394-110A-2	Sequence 1, Appli
c 29	99.4	69.5	145	2	US-07-789-917A-1	Sequence 1, Appli
c 30	99.4	69.5	145	3	US-08-702-573-4	Sequence 4, Appli
c 31	99.4	69.5	145	3	US-07-982-193-1	Sequence 1, Appli
c 32	99.4	69.5	145	3	US-09-782-378A-6	Sequence 6, Appli
c 33	99.4	69.5	165	2	US-07-989-841A-1	Sequence 1, Appli
c 34	99.4	69.5	165	2	US-08-440-738A-1	Sequence 1, Appli
c 35	99.4	69.5	165	3	US-08-471-914-1	Sequence 7, Appli
c 36	99.4	69.5	165	3	US-09-276-625-7	Sequence 7, Appli
c 37	99.4	69.5	165	3	US-10-054-665A-7	Sequence 7, Appli
c 38	99.4	69.5	165	3	US-09-782-378A-8	Sequence 8, Appli
c 39	99.4	69.5	4675	3	US-09-782-378A-1	Sequence 2, Appli
c 40	99.4	69.5	4675	3	US-10-038-972A-12	Sequence 12, Appli
c 41	99.4	69.5	4679	3	US-09-782-378A-2	Sequence 1, Appli
c 42	99.4	69.5	4680	2	US-08-254-358-1	Sequence 1, Appli
c 43	99.4	69.5	4680	2	US-08-475-391-1	Sequence 1, Appli
c 44	99.4	69.5	4680	2	US-08-709-609-1	Sequence 1, Appli
c 45	99.4	69.5	4680	6	PCT-US95-07178-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-807-802A-1
; Sequence 1, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:
US-09-807-802A-1

Query Match	100.0%	Score 143;	DB 3;	Length 4718;
Best Local Similarity	100.0%	Pred. No. 2.2e-30;		
Matches 143;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCC	60	
Db	4576	TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCC	4635	
QY	61	TGCGGACCAAGGTCCGAGACGGCAGAGCTTGCTTGTGCGGCCCCACCGAGCGAGCGA	120	
Db	4636	TGCGGACCAAGGTCCGAGACGGCAGAGCTTGCTTGTGCGGCCCCACCGAGCGAGCGA	4695	
QY	121	GCOCGACAGAGGAGTGGGCAA	143	
Db	4696	GCOCGACAGAGGAGTGGGCAA	4718	

```
RESULT 2
US-09-807-802A-19/c
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19

Query Match      84.1%; Score 120.2; DB 3; Length 4683;
Best Local Similarity 97.6%; Pred. No. 4.3e-24;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 TTGCCACCTCCCTCTCTGCGCGCTCGCTCGCTGGGGCTTGGCGACCAAGGTCGCG 78
DB 4683 TTGCCACCTCCCTCTCTGCGCGCTCGCTCGCTGGGGCTTGGCGACCAAGGTCGCG 4624

QY 79 AGACGGCAGAGCTCTCTCTGCGGCCCCCAGCGAGCGGCGCAGAGGAGTG 138
DB 4623 AGACGGCAGAGCTCTCTCTGCGGCCCCCAGCGAGCGGCGCAGAGGAGTG 4564

QY 139 GGCAG 143
DB 4563 GGCAG 4559

RESULT 3
US-09-807-802A-1/c
; Sequence 1, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:
```

US-09-807-802A-1

```
Query Match      83.2%; Score 119; DB 3; Length 4718;
Best Local Similarity 89.5%; Pred. No. 9.1e-24;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 TTACCCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCC 60
DB 143 TTACCCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCC 84

QY 61 TCGGACCAAGGTCGCGAGCGGAGCTCTCTCTGCGGCCCCCAGCGAGCGAGCGA 120
DB 83 GCGAGCAGAGCTCTGCGGCTGCGGACCTTTGCTCGCAGGCCCCACCGAGCGAGCGA 24

QY 121 GCGCGCAGAGGAGTGCGCAA 143
DB 23 GCGCGCAGAGGAGTGCGCAA 1

RESULT 4
US-09-807-802A-19
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19

Query Match      79.2%; Score 113.2; DB 3; Length 4683;
Best Local Similarity 87.3%; Pred. No. 3.6e-22;
Matches 124; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 2 TACCCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGGCGCTCGCTCGCTCGGTGGGGCC 61
DB 4542 TACCCCTAGTAGTGAGTTGCCACTCCCTCTCTATGCGCGCTCGCTCGGTGGGGCCG 4601

QY 62 GCGGACCAAGGTCGCGAGCGGAGCTCTCTGCGGCCCCCAGCGAGCGAGCGAG 121
DB 4602 GCGGAGCAGAGCTCTGCGGCTCTGCGGACCTTTGGTCCGAGCCCCACCGAGCGAGCGAG 4661

QY 122 GCGCGCAGAGGAGTGCGCAA 143
DB 4662 GCGCGCAGAGGAGTGCGCAA 4683

RESULT 5
US-09-807-802A-18
; Sequence 18, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
```



```

; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 4681
; TYPE: DNA
; ORGANISM: AAV-2
US-09-807-802A-18

Query Match          76.2%; Score 109; DB 3; Length 4681;
Best Local Similarity 85.8%; Pred. No. 5.2e-21;
Matches 121; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 3 ACCCTAGTAGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62
Db 4541 ACCCTAGTAGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62

QY 63 CGGACCAAGGTCGCGACGCGCAGAGCTCTGCTTGC CGCGGCCCCACCGAGCGAGC 122
Db 4601 GCGACCAAGGTCGCGACGCGCGCGGCTTTGCCCGCGGCTCAGTGAGCGAGC 122

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 4661 GCGCAGAGAGGAGTGGGCAA 4681

RESULT 6
US-08-525-866-1/c
; Sequence 1, Application US/08525866
; Patent No. 6207457
; GENERAL INFORMATION:
; APPLICANT: NATSOUKIS, GEORGES
; APPLICANT: FURSKY, RICHARD T.
; TITLE OF INVENTION: TARGETED NUCLEOTIDE SEQUENCE DELIVERY
; TITLE OF INVENTION: AND INTEGRATION SYSTEM
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,866
; FILING DATE: 08-SEP-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 0800-0006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-525-866-1

Query Match          75.1%; Score 107.4; DB 3; Length 145;
Best Local Similarity 85.1%; Pred. No. 9.9e-21;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTAGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62
Db 4541 ACCCTAGTAGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62

QY 5 ACCCTAGTAGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 64
Db 63 CGGACCAAGGTCGCGACGCGCAGAGCTCTGCTTGC CGCGGCCCCACCGAGCGAGC 122
Db 65 GCGACCAAGGTCGCGCGCGCGGCTTTGCCCGCGGCTCAGTGAGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 8
US-07-989-841A-1
; Sequence 1, Application US/07989841A
; Patent No. 5478745
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,841A
; FILING DATE: On even date herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-07-989-841A-1

Query Match          75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTAGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62
Db 5 ACCCTAGTAGGAGTTGCCACCTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 64

QY 63 CGGACCAAGGTCGCGACGCGCAGAGCTCTGCTTGC CGCGGCCCCACCGAGCGAGC 122
Db 65 GCGACCAAGGTCGCGCGCGCGGCTTTGCCCGCGGCTCAGTGAGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 8
```

```
US-08-440-738A-1
; Sequence 1, Application US/08440738A
; Patent No. 5869305
; GENERAL INFORMATION:
; APPLICANT: Samuleki, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,738A
; FILING DATE: May 15, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TLEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-440-738A-1

Query Match 75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 62
Db 5 ACCCTAGTGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGG 64

QY 63 CGGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCGCGCGCGCGCGAGC 122
Db 65 GCGACCAAGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 9
US-08-471-914-1
; Sequence 1, Application US/08471914A
; Patent No. 6057152
; GENERAL INFORMATION:
; APPLICANT: Samuleki, R.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM
; CURRENT APPLICATION NUMBER: US/08/471,914A
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/440,738
; EARLIER FILING DATE: 1995-05-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0

Query Match 75.1%; Score 107.4; DB 3; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 62
Db 5 ACCCTAGTGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGG 64

QY 63 CGGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCGCGCGCGCGCGAGC 122
Db 65 GCGACCAAGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 10
US-09-276-625-7
; Sequence 7, Application US/09276625
; Patent No. 6436392
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US1
; CURRENT APPLICATION NUMBER: US/09/276,625
; CURRENT FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6436392 5,478,745
US-09-276-625-7

Query Match 75.1%; Score 107.4; DB 3; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 62
Db 5 ACCCTAGTGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGG 64

QY 63 CGGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCGCGCGCGCGCGAGC 122
Db 65 GCGACCAAGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 11
US-10-054-665A-7
; Sequence 7, Application US/10054665A
; Patent No. 6897045
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
```

```
; APPLICANT: University of Iowa Research Foundation
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054,665A
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 09/276,625
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6897045 5,478,745
; US-10-054-665A-7

Query Match      75.1%; Score 107.4; DB 3; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTGGTGGGCGTG 62
   |||||
Db 5 ACCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGCTCGCTGAGGCCG 64
   |||||

QY 63 CGACCAAAAGTCCGACAGCGAGAGCTCTGCTGCGGCGCCACCGAGCGAGCGAGC 122
   |||||
Db 65 GCACCAAAAGTCCGACAGCGCGGGCTTTGCCGGGGCGCTCAGTGAGCGAGCGAGC 124
   |||||

QY 123 GCGCAGAGAGGGAGTGGGCAA 143
   |||||
Db 125 GCGCAGAGAGGGAGTGGCCAA 145
   |||||

RESULT 12
US-08-702-573-3
; Sequence 3, Application US/08702573
; Patent No. 6033885
; GENERAL INFORMATION:
; APPLICANT: LATTA, Martine
; APPLICANT: DENEFE, Patrice
; APPLICANT: VIGNE, Emmanuelle
; APPLICANT: PERRICAUDET, Michel
; TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES,
; TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,573
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/02445
; FILING DATE: 03-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00233
; FILING DATE: 28-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: 38,619
```

```
; REFERENCE/DOCKET NUMBER: ST94011-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..192
; OTHER INFORMATION: /note= "Right ITR Sequence in
; Patent No. 6033885
; OTHER INFORMATION: pXL2384"
; US-08-702-573-3

Query Match      75.1%; Score 107.4; DB 3; Length 192;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGCTGGGCGCTG 62
   |||||
Db 52 ACCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGCTCGCTGAGGCCG 111
   |||||

QY 63 CGACCAAAAGTCCGACAGCGAGAGCTCTGCTTGC CGGCGCCACCGAGCGAGCGAGC 122
   |||||
Db 112 GCGACCAAAAGTCCGCGAGCGCGGGCTTTGCCGGGGCGCTCAGTGAGCGAGCGAGC 171
   |||||

QY 123 GCGCAGAGAGGGAGTGGGCAA 143
   |||||
Db 172 GCGCAGAGAGGGAGTGGCCAA 192
   |||||

RESULT 13
US-09-276-625-6
; Sequence 6, Application US/09276625
; Patent No. 6436392
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US1
; CURRENT APPLICATION NUMBER: US/09/276,625
; CURRENT FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 272
; TYPE: DNA
; ORGANISM: AAV circular intermediate, clone pl202
; US-09-276-625-6

Query Match      75.1%; Score 107.4; DB 3; Length 272;
Best Local Similarity 85.1%; Pred. No. 1.1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGCTGGGCGCTG 62
   |||||
Db 53 ACCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGCTCGCTGAGGCCG 112
   |||||

QY 63 CGACCAAAAGTCCGACAGCGAGAGCTCTGCTGCGGCGCCACCGAGCGAGCGAGC 122
   |||||
Db 113 GCGACCAAAAGTCCGCGAGCGCGGGCTTTGGTCGCGCGGCTCAGTGAGCGAGCGAGC 172
   |||||

QY 123 GCGCAGAGAGGGAGTGGGCAA 143
   |||||
Db 173 GCGCAGAGAGGGAGTGGCCAA 193
   |||||
```

QY 63 CGGACCAAGGTCGCGAGCGGAGAGCTCTGCTGCGGCGCCACCGAGCGAGCGAGC 122
Db 4595 GCGACCAAAAGGTGCGCCGACGCGCGGGCTTTGCCGCGCGGCTCAGTGAGCGAGCGAGC 4654
QY 123 GCGCAGAGAGGAGGTGGGCAA 143
Db 4655 GCGCAGAGAGGAGGTGGCAA 4675

Search completed: November 29, 2005, 14:27:39
Job time : 20.3389 secs

RESULT 14
US-10-054-665A-6
; Sequence 6, Application US/10054665A
; Patent No. 6897045
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongshen
; APPLICANT: University of Iowa Research Foundation
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054,665A
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 09/276,625
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 272
; TYPE: DNA
; ORGANISM: AAV circular intermediate, clone p1202
US-10-054-665A-6

Query Match 75.1%; Score 107.4; DB 3; Length 272;
Best Local Similarity 85.1%; Pred. No. 1.1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 3 ACCCTAGTAGTGAGTGGCTCCCTCTCTGCGCGCTCGCTCGTGGGGCGCTG 62
Db 53 ACCCTAGTAGTGAGTGGCTCCCTCTCTGCGCGCTCGCTCGTGGGGCGCTG 112
QY 63 CGGACCAAGGTCGCGAGCGGAGAGCTCTGCTGCGGCGCCACCGAGCGAGCGAGC 122
Db 113 GCGACCAAAAGGTGCGCCGACGCGCGGGCTTTGGTCGCCCGGCTCAGTGAGCGAGCGAGC 172
QY 123 GCGCAGAGAGGAGGTGGGCAA 143
Db 173 GCGCAGAGAGGAGGTGGCAA 193

RESULT 15
US-09-782-378A-1
; Sequence 1, Application US/09782378A
; Patent No. 6916635
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 4675
; TYPE: DNA
; ORGANISM: Human adeno-associated virus 2
US-09-782-378A-1

Query Match 75.1%; Score 107.4; DB 3; Length 4675;
Best Local Similarity 85.1%; Pred. No. 1.4e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 3 ACCCTAGTAGTGAGTGGCTCCCTCTCTGCGCGCTCGCTCGTGGGGCGCTG 62
Db 4535 ACCCTAGTAGTGAGTGGCTCCCTCTCTGCGCGCTCGCTCGTGGGGCGCTG 4594